AD-A055 613 TEXAS TECH UNIV LUBBOCK INST FOR ELECTRONICS SCIENCE F/G 12/1
RESOLUTION SPACE, NETWORKS, AND NON-SELF-ADJOINT SPECTRAL THEOR--ETC(U)
1978 R SAEKS, L TUNG, R A DECARLO AFOSR-74-2631
UNCLASSIFIED AFOSR-TR-78-1045 NL

1 OF 2
AD
A055 613

AFOSR-TR- 78-1045

FOR FURTHER TRAN

A055434

2

AD A055613

Resolution Space, Networks, and Non-Self-Adjoint Spectral Theory II

by

R. Saeks, L. Tung, R. DeCarlo, M. Strauss
R. M. DeSantis, D. Brandon

This research supported in part by
AFOSR Grant 74-2631

AD No. ____
DDC FILE COPY

DDC
JUN 23 1978
RECEIVED
E

Institute for Electronics Science

TEXAS TECH UNIVERSITY
Lubbock, Texas 79409

78 06 19 100

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 78-1045	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) RESOLUTION SPACE, NETWORKS, AND NON-SELF-ADJOINT SPECTRAL THEORY II.		5. TYPE OF REPORT & PERIOD COVERED Interim rept.,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. Saeks, L. Tung, R.A. DeCarlo, M. Strauss, R.M. DeSantis D. Brandon		8. CONTRACT OR GRANT NUMBER(s) AFOSR-74-2631
9. PERFORMING ORGANIZATION NAME AND ADDRESS Texas Tech University, Department of Electrical Engineering, Lubbock, Texas 79409		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A6 A6
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332		12. REPORT DATE 1978
		13. NUMBER OF PAGES 142
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 139 p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Resolution space, stability theory, Nyquist criterion, Wiener-Hopf filtering, relativistic systems.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes research done in which the theory of operators defined on a Hilbert resolution space is applied to problems arising in system theory. Specific topics include Wiener-Hopf filtering, generalizations of the Nyquist stability criterion to nonlinear and time-varying systems, and a formulation of the theory of systems subjected to relativistic effects.

410 354

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

RESOLUTION SPACE, NETWORKS, AND

NON-SELF-ADJOINT SPECTRAL THEORY II*

R. Saeks, L. Tung, R.A. DeCarlo, M. Strauss

R.M. DeSantis, and R. Saeks

1978

*This research supported in part by AFOSR gran 74-2631.

CONTENTS

WIENER-HOPF FILTERING IN HILBERT RESOLUTION SPACE*

L. Tung, R. Sakes and R.M. DeSantis

*This research supported in part by AFOSR grant 74-2631 and CNRC grant 8244.

I. Introduction

The development of the theory of causal operators defined on a resolution space was initiated a decade ago in response to the failure of classical Hilbert space methods to yield a solution to the quadratic optimization problems of mathematical system theory. Although this theory has now achieved a considerable degree of maturity yielding viable solutions to problems arising in network synthesis [1], feedback system stability [9], sensitivity theory [1], and stochastic processes [7,8], the solution to the original quadratic optimization problems has remained elusive. Porter and DeSantis [12,23] have solved a deterministic servomechanism problem, Stienberger, Silverman, and Schumitzky [22], have solved a deterministic regulator problem and Saeks [7], has solved a stochastic identification problem. The original goal of a general theory for quadratic optimization, however, has yet to be achieved.

The purpose of the present paper is to present a derivation of the Wiener-Hopf filter using resolution space techniques. Although still restricted, we believe that the tools employed are indicative of the techniques which will eventually lead to a general theory of quadratic optimization. Indeed, even for this restricted filter the deriviation requires several recently developed results from the theory of causal operators. These include:

i) the additive decomposition theorem for Hilbert-Schmidt operators [10],

ii) the miniphase factorization theorm [7, 19],

iii) the theory of resolution space valued stochastic processes [7, 8, 11],

iv) the quasi-nilpotence theorem for strictly causal operators [9, 10].

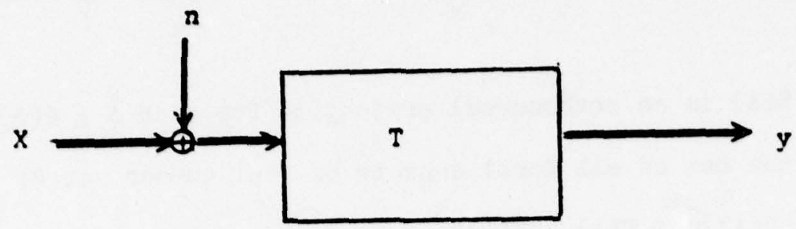


Figure 1. Wiener-Hopf Filter

A basic filtering problem is illustrated in Figure 1, where the signal "X" and noise "n" are mixed together. The goal is to pass the mixed signal through a filter T (to be designed) to get an output y such that y is the "best copy" of X that can be achieved. When X and n are assumed to be stationary, zero-mean and independent stochastic processes and the "best copy" of X is defined in the sense that the error e (= X - y) has a minimum variance, the solution is nothing but the classical Wiener-Hopf filter. [2] In this paper, we would like to formulate this problem in Hilbert resolution spaces, i.e. we assume X and n to be Hilbert space valued random variables (zero-mean and independent) and attempt to find a filter T (causal operator) such that the output y is the best copy of X in the sense that the "variance operator" of the error is minimal.

Before we continue, let us explain the terminology we have just used.

Hilbert Resolution Space

By a Hilbert Resolution Space, we mean a 2-tuple, (H,E), where H is a Hilbert space over the real field R and E is a "so-called" resolution of identity (or spectral measure) in H. [1] A resolution of identity E is a family of bounded linear operators, $E(\Delta)$, on H defined for each Borel subset Δ of the real number set R, satisfying the following conditions:

i) $E(\Delta)$ is an orthonormal projection for each $\Delta \in \mathcal{B}(R)$ = the set of all Borel subsets of real number set R; i.e. $[E(\Delta)]^2 = E(\Delta) = E(\Delta)^*$

ii) $E(\Delta_1)\, E(\Delta_2) = E\,(\Delta_1 \cap \Delta_2),\ \forall\ \Delta_1, \Delta_2 \in \mathcal{B}(R)$

iii) $E^t \overset{def}{=} E((-\infty, t))$ is continuous in the strong operator topology, goes to 0 as T goes to $-\infty$, and goes to I_H (identity mapping on H) as T goes to ∞.

Defined as an operator from (H, E) to $(\hat{H}, \hat{E})$, T is said to be causal if

$$E^t X_1 = E^t X_2 \Longrightarrow \hat{E}^t T X_1 = \hat{E}^t T X_2$$

A special class of causal operators termed as left-miniphase can be defined by the following condition

$$E^t X_1 = E^t X_2 \Longleftrightarrow \hat{E}^t T X_1 = \hat{E}^t T X_2$$

The significance of being left-miniphase is that the inverse operator is also causal when the invertibility of the operator is guaranteed. For a more detailed survey of the properties of a Hilbert resolution space, the reader is referred to [1, 14, 20].

Hilbert Resolution Space Valued Random Variables

When talking about random variables in a Hilbert space, the structure of the resolution of identity is redundent. But for the purpose of this paper, "resolution" is included in the title to avoid ambiguity. Let X and n be random variables [4] taking values in a Hilbert space H with appropriate probability measure. X is said to have finite first moment if for each $h \in H$

$$E\,|(X,h)| < \infty \tag{A}$$

and $E(X,h)$ is continuous in h.

Here "E" denotes the expected value of a scalar valued random variable with respect to the probability space underlying X. A random variable satisfying condition (A) has a "mean" M_X which is defined by the equation:

$$E(X,h) = (h, M_X).$$

The existance and the uniqueness of M_X is guaranteed by condition (A) following from the Riesz representation theorem [21].

In the sequel we deal with zero-mean random variables unless otherwise specified. Such a random variable is said to have finite second moment if

$$E(X,h)^2 < \infty \text{ for all } h \in H \tag{B}$$

and it is continuous in h.

As such, X has a "variance operator" Q_X which is defined as follows

$$E(X,h)(X,g) = (h, Q_X g) \quad \forall\, h\, g \in H$$

The existence and uniqueness of Q_X is assured by the representation theorem for bilinear functionals on Hilbert space [8] together with condition (B). More generally, given two zero-mean processes X and n with finite second moment, a covariance operator Q_{Xn} can be defined as follows

$$E(X,h)(n,g) = (h, Q_{Xn}\, g)\ ,\ \forall\, h \in H$$

clearly $Q_X = Q_{XX}$.

The covariance operators satisfy the following:

i) $Q_{(LX)(Mn)} = L\, Q_{Xn}\, M^*$, for bounded linear operators L and M

ii) $Q_{X+n} = Q_X + Q_{Xn} + Q_{nX} + Q_n$

iii) $Q_{Xn} = Q^*_{nX}$, in particular, $Q_X = Q^*_X$

iv) Q_X is positive, i.e. $(h, Q_X h) > 0$

$\forall\, h \in H$

v) $E||X||^2 < \infty$ if and only if Q_X is nuclear in which case

$E||X||^2 = Tr[Q_X]$

Finally, we say X and n are independent if $Q_{Xn} = 0$.

II. The Optimal Filter

With the terminology defined above, we can now formally state the problem as follows:

Let X and n be Hilbert Resolution space $(H,\mathcal{E})$ valued random variables. X and n are zero-mean and independent. X and n have Q_X and Q_n as their variance operator, respectively. We want to find a causal filter operator T on $(H,\mathcal{E})$ such that the output y of the filter (with X + n as the input) is the best copy of X in the sense that the error e (defined as X - y) has a minimal variance operator Q_e. Note here that Q_e is a positive operator hence it makes sense to talk about the minimal Q_e in the partial ordering of positive operators.

Even though the problem is stated above in its most general form, it is not solved for the general case. However, we do find an optimal filter for a special case, characterized by the following assumptions:

i) Q_X is Hilbert-Schmidt.

ii) $Q_X + Q_n$ is invertible.

iii) The optimal filter T to be found is restricted to the class of Hilbert-Schmidt operators. Note: the assumptions are compatible with the interpretation of X as a random signal and n as a noise process.

With these assumptions in mind, let us derive the variance operator Q_e.

Since $e \overset{\text{def}}{=} X - y$,

$$e = X - y = X - T(X + n) = (I - T)X + Tn$$

Hence

$$\begin{aligned} Q_e &= Q_{(I-T)X} + Q_{Tn} \\ &= (I-T)\, Q_X\, (I-T)^* + T\, Q_n T^* \\ &= T\,(Q_X + Q_n)\, T^* - TQ_X - Q_X T^* + Q_X. \end{aligned}$$

In deriving Q_e, independence of X and n has been employed. We have also noted that $Q_{X+n} = Q_X + Q_n$ is positive and self-adjoint. Following the operator factorization theorem [7, 19, 20], there exists a Hilbert resolution space* $(\hat{H}, \hat{E})$ and a linear operator F from $(\hat{H}, \hat{E})$ to (H, E) such that (a) F is left-miniphase and (b) $Q_X + Q_n = FF^*$. Furthermore, F is invertible since $Q_X + Q_n$ is invertible. Therefore F^{-1} is causal. Then we can rewrite Q_e as follows:

$$\begin{aligned} Q_e &= T\,F\,F^*\,T - TQ_X - Q_X\,T^* + Q_X \\ &= [TF - Q_X(F^*)^{-1}](F^*\,T^* - F^{-1}Q_X) + Q_X - Q_X F^{*-1}\,F^{-1}\,Q_X \\ &= [TF - Q_X(F^*)^{-1}]\,[TF - Q_X(F^*)^{-1}]^* + Q_X - Q_X\,(F\,F^*)^{-1}\,Q_X. \end{aligned}$$

In the above equation, Q_X and $Q_X(F\,F^*)^{-1}\,Q_X$ are both positive and independent of T. Therefore, finding the minimum of Q_e is the same as finding the minimum of $[TF - Q_X\,(F^*)^{-1}]\,[TF - Q_X(F^*)^{-1}]^*$ denoted as Q(T) from now on.

Q(T) is in quadratic form. The minimum Q(T) occurs when $TF = Q_X(F^*)^{-1}$. To fulfill this equation, we need a filter T which is equal to $Q_X(F^*)^{-1}\,F^{-1} = Q_X(F\,F^*)^{-1} = Q_X(Q_X + Q_n)^{-1}$. Unfortunately, this operator is not necessarily causal. In order to find an optimal causal filter, we decompose $Q_X(F^*)^{-1}$ into two terms;

$Q_X(F^*)^{-1} = A + C$, where A is the strictly anti-causal part of $Q_X(F^*)^{-1}$ and C is the causal part. The existance and the uniqueness of

*This space has been shown to be the reproducing kernel resolution space for $Q_X + Q_n$ [7,19], which does not, in general, coincide with $(\hat{H},\hat{E})$. Indeed, the factorization may not exist if one requires that (H,E) and $(\hat{H},\hat{E})$ coincide.

the decomposition are guaranteed by the fact that $Q_X(F^*)^{-1}$ is Hilbert-Schmidt. Readers are referred to [1, 7, 9, 10, 19, 20] for the terminology and details.

Substituting $Q_X(F^*)^{-1}$ back into the equation for Q(T), we obtain

$$Q(T) = (T\,F - C - A)\,(T\,F - C - A)^*$$
$$= (T\,F - C)\,(T\,F - C)^* - (T\,F - C)\,A^* - A\,(T\,F - C) + A\,A^*$$

In this equation, $A(T\,F - C)^*$ is strictly anti-causal and $(T\,F - C)\,A^*$ is strictly causal.

By the assumption that T and Q_X are Hilbert-Schmidt, each term in the equation for Q(T) is found to be nuclear. Therefore, the trace of Q(T) is taken.

$$Tr[Q(T)] = Tr\,[(T\,F - C)\,(T\,F - C)^*] + Tr\,[A\,A^*]$$

Two terms are dropped in the above equation due to the fact that the trace of any strictly causal (or anti-causal) nuclear operator is zero [9, 10]. It is also known that the trace of any positive nuclear operator is positive. Hence, the minimum of Tr[Q(T)] occurs when T F = C, i.e. when we have a filter $T = C\,F^{-1}$. This filter $C\,F^{-1}$ is causal following our derivation. Now the only problem left is to verify that $Q(C\,F^{-1})$ is minimal. The verification is straigtforward following from the fact that a positive, self-adjoint operator is zero if and only if its trace is zero [20].

To summarize what we have done in this section, we formulate the following theorem.

Thm. i) Let X and n be Hilbert Resolution Space (H, E) valued random variables.

ii) X and n are zero-mean and independent.

Q_X and Q_n denote the variance operators of X and n, respectively.

iii) Q_X is assumed to be Hilbert-Schmidt and $Q_n + Q_X$ is invertible.

iv) F is a left-miniphase factorization of $Q_X + Q_n$.

The, the optimal causal filter operator among the class of the Hilbert-Schmidt operators is $C\,F^{-1}$, where C is the causal part of the operator $Q_x(F^*)^{-1}$.

III. Conclusions

The Wiener-Hopf filter derived above is sketched diagramatically in Figure 2. The first transformation, denoted by F^{-1}, is usually termed the whitening filter since its output, the innovations process, is white noise; i.e. it has a memoryless covariance operator [7]. It is significant that for the filter to be well defined it was necessary to take $(\hat{H},\hat{E})$ to be the reproducing kernel resolution space for $Q_x + Q_n$ rather than the given resolution space (H,E). As such, one may conclude that the innovations process naturally "lives" in this reproducing kernel space thus yielding a further justification for the study of this abstract resolution space even though the given system is defined on a concrete resolution space.

$$\begin{tikzcd}
 & (\hat{H},\hat{E}) \arrow[dr, "C"] & \\
(H,E) \arrow[ur, "F^{-1}"] \arrow[rr, "F^{-1}C = T"'] & & (H,E)
\end{tikzcd}$$

Figure 2. Diagramatic Representation of the Wiener-Hopf Filter

Although the above derivation is restricted to a very specialized filtering problem we believe that the techniques employed are indicative of those which will eventually lead to a general theory of quadratic optimization in resolution space. Indeed, the authors have already made significant progress towards the generalization of the above concepts to stochastic control and estimation problems for systems described by both input-output and state models [20,24]. Of course, these results, as with the above Wiener-Hopf filter, apply to distributed and time-variable systems as well as the classical LLF systems.

A careful inspection of the derivation of the above filter will reveal that the restriction that $Q_x + Q_n$ be invertible can be dropped by exploiting the fact that the miniphase factor of $Q_x + Q_n$ is one-to-one for arbitrary covariance operators [7, 19]. As such, if one replaces F^{-1} by F^{-L} (left inverse) the above derivation may be carried out without the invertibility assumption. Finally, we note that the Hilbert-Schmidt assumptions are required only to make the trace well defined. The derivation, however, remains (formally) valid if one allows arbitrary operators and hence infinite values for the trace. Alternatively, one may replace the trace by the memoryless part transformator [1, 10] obtaining the Wiener-Hopf filter by minimizing the memoryless part of the error covariance (in the partial ordering of positive operators). This results is an optimal (though non-unique) filter whenever the operator $[Q_x(F^*)^{-R}]$ has a well defined additive decomposition into causal and strictly anti-causal parts.

IV. REFERENCES

1. Saeks, R., Resolution Space, Operators and Systems, New York, Springer-Verlag, 1973.

2. Cooper, G.R. and Mcgillem, C.D., Probabilistic Methods of Signal and System Analysis, New York, Holt, Rinehart and Winston, Inc., 1971.

3. Parthasarathy, K.R., Probability Measures on Metric Spaces, Academic Press, 1967.

4. Balakrishnan, A.V., Introduction to Optimization Theory in a Hilbert Space, New York, Springer-Verlag, 1971.

5. Grenander, U., "Probabilities on Algebraic Structures", Wieley, New York, 1963.

6. Luenberger, D.G., Optimization by Vector Space Methods, New York, J. Wiley and Sons, 1969.

7. Saeks, R.,"Reproducing Kernel Resolution Space and its Applications,"in Journal of the Franklin Institute, Vol. 302, No. 4, pp. 331-355, Oct. 1976.

8. Kailath, T. and Duttweiler, D., "An RKHS Approach to Detection and Estimation Problems-Part III: Generalized Innovations and a Likelihood-ratio Formula", IEEE Trans. Info. Thy., Vol. IT-18, pp. 730-745, 1972.

9. DeSantis, R.M., "Causality, Strict Causality and Invertibility in Hilbert Resolution Spaces", SIAM J. Control, Vol. 12, No. 3, Aug. 1974.

10. DeSantis, R.M., "Causality Structure of Engineering Systems", Ph.D. Thesis, University of Michigan, Sept. 1971.

11. Duttweitler, E., "Reproducing Kernel Hilbert Space Techniques for Detection and Estimation Problems", Ph.D. Thesis, Stanford University, 1970.

12. Porter, W.A., "A Basic Optimization Problem in Linear Systems", Math. Sys. Thy., Vol. 5, pp. 20-44, (1971).

13. Aronazajn, N., "Theory of Reproducing Kernels", Trans. on the AMS, Vol. 63, pp. 337-404, 1950.

14. Porter, W.A., "Some Circuit Theory Concepts Revisited", Int. Journ. on Control, Vol. 12, pp. 443-448, 1970.

15. Schnure, W.K., "Controllability and Observability Criteria for Causal Operators on a Hilbert Resolution Space", Proc. of the 14th MSCT, Univ. of Denver, 1971.

16. Gohberg, I.C. and Krein, M.G., Introduction to the Theory of Linear Non-self-adjoint Operators, American Mathematical Society, 1969.

17. Youla, D.C., Jabr. H.A. and Bongiorno, J.J., Jr., "Modern Wiener-Hopf Design of Optimal Controllers-Part II: The Multivariable Case", IEEE Trans. of Automatic Control, Vol. AC-21, No. 3, June 1976.

18. Yosida, K., Functional Analysis, Springer, Heidelberg, 1966.

19. Saeks, R., "The Factorization Problem - A Survey", Proc. of the IEEE, Vol. 64, No. 1, Jan. 1976.

20. Tung, L., "Random Variables, W-H Filtering and Control Formulated in Abstract Spaces", Ph.D. Dissertation, Texas Tech University, Lubbock, Tx., Aug. 1977.

21. Akhiezer, N.I. and Glazman, I.M., Theory of Linear Operators on Hilbert Space, Unger, New York, 1963.

22. DeSantis, R.M., "Causality Theory in Systems Analysis", Proc. of IEEE, Vol. 64, No. 1, Jan. 1976.

23. Steinberger, M., Schumitzky, A. and Silverman, L., "Optimal Causal Feedback Control of Linear Infinite - Dimensional Systems", Unpublished notes, Univ. of Southern Calif.

24. DeSantis, R.M., Saeks, R. and Tung, L., "Basic Optimal Estimation and Control Problems in Hilbert Space", Unpublished notes, Texas Tech Univ.

WIENER-HOPF TECHNIQUES IN RESOLUTION SPACE*

L. Tung and R. Saeks

*This research supported in part by AFOSR grant 74-2631.

WIENER-HOPF TECHNIQUES IN RESOLUTION SPACE

L. Tung and R. Saeks
Dept. of Electrical Engineering
Texas Tech University
Lubbock, Texas 79409

INTRODUCTION

Wiener-Hopf filtering is a widely used technique in certain kinds of optimization problems. The purpose of this paper is to formulate Wiener-Hopf filtering in abstract spaces (reflexive Banach resolution spaces) and to examine problems involved for the formulation and the solving of the Wiener-Hopf filter.

Referring to what has been done in the frequency domain of the classical Wiener-Hopf filtering[1], we've found five major problems for the formulation of Wiener-Hopf filtering in abstract spaces. They are

i. Random variables in abstrace spaces

ii. Causality

iii. Operator factorization

iv. Operator decomposition

v. Optimization.

These problems are briefly introduced as follows:

i. Random process can be thought of as a random variable which takes values in a function space. In order to do so, we need an adequate probability measure over the space involved. Fortunately, this kind of probability measure has been defined over metric space[2]. For our purposes, we assume that the space involved is reflexive Banach space, not only because this kind of space possesses nice properties but also because stochastic concepts such as "mean" and "variance operation" can be defined therein. Random variables taking values in reflexive Banach space is discussed in section II with probability measure assumed implicitly.

ii. Concepts of causality have been introduced into Hilbert space-the so-called Hilbert resolution space[3]. In section III, we extend the works done for Hilbert space to Banach space. Concepts of causality, such as causal, anti-causal, miniphase and maxiphase, are defined. Emphases are given to reflexive Banach resolution space.

iii. Operators to be factorized in the form of KK*, where K* denotes the adjoint of K, have to be "positive" and "self-adjoint". These commonly-used properties among operators on Hilbert space can be extended to operators which map reflexive Banach space to its dual space. Factorization theorem is given in section IV. Factor operator K is required to be left-miniphase.

iv. The decomposition of operators over Hilbert space is treaded in Ref. 3. For operators over Banach spaces, this problem is still under research. For our convenience, operators are restricted to those which guarantee the decomposition.

v. As in the classical Wiener-Hopf filtering, we would like to minimize the variance of the error. However, when Wiener-Hopf filtering is formulated in reflexive Banach space, the variance of the error is a positive and self-adjoint operator which can only be minimized in the partial ordering of the positive operators. This subject is treaded in section V.

BANACH SPACE VALUED RANDOM VARIABLES

The theory of Banach space valued random variables has been studied in Ref. 2. For our purpose, we discuss reflexive Banach space valued random variables with probability measure over the space assumed implicitly.

The development follows that of Parthasarathy (2) and Balakrishnan (4); the reader is referred to these works for the details.

Let ρ, π denote finitely additive random variables taking values in a reflexive Banach space B. For such random variables, we assume

$$E\{|(\rho, x^*)|\} < \infty \text{ , for all } x^* \in B^* \qquad (2.1)$$
$$E\{(\rho, x^*)\} \text{ is continuous in } x^*$$

Here $E\{\cdot\}$ denotes the expected value of a scalar valued random variable with respect to the probability space underlying ρ. For random variables which satisfy condition (2.1), there is a unique vector m_ρ in B satisfying

$E\{(\rho, x^*)\} = (m_\rho, x^*)$, for $x^* \in B^*$.

m_ρ is termed as the mean of random variable ρ. As in most stochastic processes, mean is not our prime concern. Therefore, in the sequel we only deal with zero-mean random variables. For such random variables, we further assume

$$E\{|(\rho, x^*)\ (\pi, y^*)|\} < \infty,$$
$$\text{for all } x^*, y^* \in B^* \qquad (2.2)$$
$$E\{(\rho, x^*)\ (\pi, y^*)\}$$
$$\text{is continuous in } x^* \text{ and } 6^*$$

It can be shown that condition (2.2) implies condition (2.1). Now let's take a look at $E\{(\rho, x^*)\ (\pi, y^*)\}$. If we fix y^*, then $E\{(\rho, x^*)\ (\pi, y^*)\}$ is a bounded linear functional on B^* (so an element of $B^{**}=B$). This means that there exists a unique p_{y^*} in B such that $E\{(\rho, x^* (\pi, y^*)\}$ $=(p_{y^*}, x^*)$ for $x^* \in B^*$. Define a mapping $Q_{\rho\pi} = B^* \to B$, by $Q_{\rho\pi} y^* = p_{y^*}$. Hence $E\{(\rho, x^*)\ (\pi, y^*)\} = (Q_{\rho\pi}\ y^*, x^*)$.

It can be easily proven that $Q_{\rho\pi}$ is linear. Moreover, $Q_{\rho\pi}$ is bounded. $Q_{\rho\pi}$ is termed as the covariance operator of random variables ρ and π. Covariance operators satisfy following conditions:

i. $Q_{(L\rho)\,(M\pi)} = L\, Q_{\rho\pi} M^*$, where L and M are linear bounded operators on B.

ii. Define $Q_\rho = Q_{\rho\rho}$, then $Q_{\rho+\pi} = Q_\rho + Q_{\rho\pi} + Q_{\pi\rho} + Q_\pi$. Q_ρ is called the variance operator of ρ.

iii. $Q_{\rho\pi} = Q^*_{\pi\rho}$, in particular $Q_\rho = Q_\rho{}^*$

iv. Q_ρ is positive in the sense that $(Q_\rho\, y^*, y^*) = E\{(\rho, y^*)^2\} \geq 0$, for all $y^* \in B$.

These conditions result from straight forward manipulation of the defining equation for the covariance operator. Using $Q_{\rho\pi}$, we say that ρ and π are independent if $Q_{\rho\pi} = 0$.

BANACH RESOLUTION SPACE

By a Banach resolution space, we mean a 2-tuple, $(B, {}_BF)$, where B is a Banach space and ${}_BF$ is the so-called resolution of identity in B, which is defined in the following:

(A) Resolution of identity

Definition 3.1. Let B be a Banach space. By a resolution of identity, ${}_BF$, in B, we mean a family of linear bounded operators, ${}_BF(\Delta)$, on B defined for each Borel subset, Δ, of the real number set R, satisfying the followings:

i. ${}_BF(R) = I_B$-identity operator on B

ii. ${}_BF(\Delta_1) \cdot {}_BF(\Delta_2) = {}_BF(\Delta_1 \cap \Delta_2)$, for all Δ_1, $\Delta_2 \in \beta(R)$ - the set of all Borel subsets of R.

iii. ${}_BF(\bigcup_1^n \Delta_i) = \sum_1^n {}_BF(\Delta_i)$, where $\{\Delta_i\}_1^n$ is a finite set of disjoint Borel subsets of R.

iv. $||{}_BF(\Delta)\, x\, || \leq ||x||$, for all $\Delta \in \beta\,(R)$ and $x \in B$ (Equivalent statement: Norm of ${}_BF\,(\Delta)$ is either 0 or 1)

The subscript on the left in the notation, ${}_BF$, is to notify that the resolution of identity is defined over space B and will be dropped if no ambiguity would result.

Working with a Banach resolution space, $(B, {}_BF)$, it is natural and important to ask whether we can define a resolution of identity in B*, the dual space of B. The following theorem gives us the answer.

Theorem 3.1. Let $(B, {}_BF)$ be a Banach resolution space, then $\{{}_BF^*(\Delta)\ |\Delta \in \beta\,(\Delta)\}$ is a resolution of identity in B*-termed as the induced resolution space, $(B^*, {}_BF^*)$.

With the resolution of identity defined as above, we'd like to point out that although Hilbert space is a special case of Banach space, Definition 3.1 does not lead to a Hilbert resolution space[3]. In Hilbert resolution space, the resolution of identity, $\{E(\Delta)\ |\Delta \in \beta\,(R)\}$, satisfies an additional condition, i.e. $E^*(\Delta) = E(\Delta)$.

Example. Let p, q $\in$ R, such that $1/p + 1/q = 1$. Then L_p is a reflexive Banach space with dual space L_q. For each $f \in L_p$, define $(f,q) = \int_{-\infty}^{\infty} f(t)\, g(t)\, dt$. Let $F(\Delta)\, f(t) = \chi(\Delta)\, f(t) = \begin{cases} 0, & t \notin \Delta \\ f(t), & t \in \Delta \end{cases}$

It is easy to show that $\{F(\Delta)\ |\ \Delta \in \beta(R)$ is a resolution of identity in L_p and $F^*(\Delta) = \chi(\Delta)$ for all $\Delta \in \beta\,(R)$.

(B) Concepts of causality

Definition 3.2. Let $(X, {}_XF)$, $(Y, {}_YF)$ be Banach resolution spaces. T: $X \to Y$, is a linear bounded operator

(i) T is causal if ${}_XF^t x_1 = {}_XF^t x_2 \to {}_YF^t T x^1 = {}_YF^t T x_2$, where ${}_BF^t = {}_BF(-\infty, t)$, B= X, y.

(ii) T is anti-causal, if $X^F t\, x_1 = X^F t\, x_2 \to T^F t\, T x_1 = {}_YF^t T x_2$, where ${}_BF_t = {}_BF\,(t,\infty)$, B = X, Y.

(iii) T is memoryless, if T is causal and anti-causal.

(iv) T is left-miniphase if ${}_XF^t x_1 = {}_XF^t x_2 \Longleftrightarrow {}_YF^t T x_1 = {}_YF^t T x_2$

(v) T is left maxiphase if ${}_XF_t\, x_1 = {}_XF_t\, x_2 \Longleftrightarrow {}_YF_t\, T x_1 = {}_YF_t\, T x_2$

(vi) T is right-miniphase, if $\overline{T\, {}_XF_t \{X\}} = {}_YF_t \{Y\}$

(vii) T is right-maxiphase, if $\overline{T\, {}_XF^t \{X\}} = {}_YF^t \{Y\}$

According to above definitions, we've found the following results:

(1) Miniphase, left-or-right-, implies causality.

(2) Maxiphase, left-or-right-, implies anti-causality.

(3) When X and Y are reflexive, we have

(a) T is causal $\Longleftrightarrow$ T* is anti causal

(b) T is left-miniphase $\Longleftrightarrow$ T* is right-maxiphase.

(4) When X and Y are reflexive and T is invertable, we have

(a) Miniphases are equivalent, so are maxiphases.

(b) T is miniphase $\Rightarrow$ T and T^{-1} causal

T is maxiphase $\Rightarrow$ T and T^{-1} anti-causal.

Readers are referred to Ref. 5 for the details.

OPERATOR FACTORIZATION

Not every operator over arbitrary Banach spaces can be factorized in the form desired. For our purpose the desired form of factorization is K K*.

An operator to be factorized in this form has to be positive and self-adjoint. These two commonly-used properties for operators over Hilbert space can be extended to operators that map from reflexive Banach space to its dual space. They are defined as follows:

Definition 4.1.

(i) Let B be a reflexive Banach space.

(ii) $Q= B \to B^*$, is linear and bounded. Q is said to be positive if $(x, Qx) \geq 0$, for each $x \in B$. Q is said to be self-adjoint if $Q^* = Q$. Note that $Q^* : B^{**}=B \to B^*$ so it makes sense to compare Q with Q^*.

For positive and self-adjoint operators, we have the following theorem:

Theorem 4.1.

(i) Let B be a reflexive Banach space.

(ii) $Q : B \to B^*$, is linear, bounded, positive and self-adjoint.

Then there exist a Hilbert space H and a linear bounded operator $K : H \to B$, such that $Q = K K^*$[6].

When dealing with Banach resolution spaces, the usefulness of operator factorization is limited unless the factor operator possesses certain causal properties. Referring to factorization of the spectral density in classical Wiener-Hopf filtering, we have found what we need is a factorization theorem which gives a causal operator and guarantees a causal inverse once the existance of a inverse is granted, i.e. a theorem that gives a miniphase factorization. Based on Theorem 4.1, we construct the resolution of identities in spaces involved and we come up with the following theorem.

Theorem 4.2.

(i) Let (B,F) be a reflexive Banach resolution space. (B^*, F^*) denotes the induced resolution space.

(ii) $Q = (B,F) \to (B^*, F^*)$, is linear, bounded, positive and self-adjoint. Then there exist a Hilbert resolution space (H,E) and a linear bounded operator $K = (H, E) \to (B^*, F^*)$ such that

1. $Q = K K^*$
2. K is a left-miniphase
3. The factorization is unique up to a memoryless unitary transformation.

For the proof of this theory, please refer to Ref. 5.

WIENER-HOPF FILTERING FORMULATED IN REFLEXIVE BANACH SPACE

With the preparation of sections II, III and IV, now we are ready for the formulation of Wiener-Hopf filtering. The formulation is done as follows:

Let X, n be random variables taking values in a reflexive Banach resolution space (B, F). X denotes the signal and n the noise. Both X and n satisfy condition (2.2) in section II and they are assumed to be zero-mean and independent. As such, X and n have Q_X and Q_n as their variance operators respectively. The problem we are facing is to find a filter, $T : B \to B$, linear and causal, to operate on $X + n$ such that the error, defined as $x-y$, where y is the output of T, has a variance operator that is minimal in the partial ordering of the positive operators. We will describe this ordering right after we find the variance of the error.

Let e denote the error and Q_e denote its variance operator. Since $e \overset{def}{=} X - y = X - T(X+n) = (I-T)X + Tn$ we have

$Q_e = (I-T)\, Q_X\, (I-T)^* + T\, Q_n\, T^*$, following from the results in section II. Rearranging terms in Q_e, we get

$$Q_e = T(Q_X+Q_n)T^* - Q_X\ T^* - T\ Q_X + Q_X.$$

Q_e is dependent on T. We write Q_e (T) to notify the dependence. Q_e (t_o) is said to be minimal for some filter T_o, if

$$Q_e(T) \leq Q_e(T_o) \Rightarrow Q_e(T) = Q_e(T_o) \quad (A \leq B \text{ if } (B-A) \text{ is positive}).$$

In the equation for Q_e, $Q_X + Q_n$ represents the variance operator of X+n, hence is positive and self-adjoint. Therefore, by Theorem 4.2, there exists a resolution space (H,E) and a linear bounded operator $K = (H,E) \rightarrow (B,F)$, such that (a) $Q_e = K\ K^*$ (b) K is a left-miniphase. Without further assumptions, the formulation would be stuck right here. At this point, what we need is an invertable factor operator K. The invertability of K can be guaranteed by the invertability of $Q_X + Q_n$. There are several ways to secure the invertability of $Q_X + Q_n$. One way is to assume that $Q_X + Q_n$ in onto and Q_n is positive definite. With an invertable factor K, Q_e can be rewritten an

$$Q_e = T\ K\ K^*\ T^* - Q_X\ T^* - K\ Q_X + Q_X = \{TK-Q_X(K^*)^{-1}\}\ \{K^*T^*-T^{-1}Q_X\} +$$
$$Q_X - Q_X(K^*)^{-1}\ K^{-1}\ Q_X = \{TK-Q_X(K^*)^{-1}\}\ \{TK-Q_X(K^*)^{-1}\}^* + Q_X - Q_X\ (KK^*)^{-1}Q_X$$

The last two terms in the above equation, Q_X and $Q_X(KK^*)^{-1}Q_X$, are positive and independent of T. Hence, to find the minimal of Q_e is the same as to find the minimal of $\{TK-Q_X(K^*)^{-1}\}\ \{TK-Q_X\ (K^*)^{-1}\}^*$ - denoted as Q(T) in the sequel. Right now, we are facing the same kind of problem as in classical Wiener-Hopf filtering. Minimal Q(t) occurs when $T = Q_X(K^*)^{-1}K^{-1}$, but it does not represent a causal system in general. In order to get a possible optimal causal filter, can we decompose $Q_X(K^*)^{-1}$ into "causal part" and "strictly anti-causal" (a term to be generalized in Banach resolution space) and under what conditions can we do so? This subject has been treated in Ref. 3 and Ref. 7 when the reflexive Banach space happens to be a Hilbert space. However in reflexive Banach resolution spaces, the subject is still under research. While we follow the same pattern as that of classical Wiener-Hopf filtering in frequency domain, we would like to ask whether this decomposition would work and how it would. The same question in classical Wiener-Hopf filtering is not directly answered in frequency domain. In order to find the answer, let's assume the decomposition. Let

$Q_X(K^*)^{-1} = C + A$, where C is the causal part of $Q_X(K^*)^{-1}$ and A is the "strictly anti-causal part" (a term to be generalized in Banach resolution space). Then

$$\begin{aligned} Q(T) &= \{TK-C-A\}\{TK-C-A\}^* \\ &= \{TK-C\}\{TK-C\}^* - A\{TK-C\}^* \\ &= \{TK-C\}A^* + AA^* \end{aligned}$$

To claim TK-C=0 is the condition for minimal A(T), we should demonstrate that those cross terms, $\{TK-C\}A^*$ and $A\{TK-C\}^*$, have no effect on the ordering of Q(T). Again when the reflexive Banach resolution space is Hilbert resolution space, we've found two ways to achieve this. The first one is to take the trace of Q(T). Surely, work has to be done to guarantee Q(T) being nuclear. The second one is to take the memoryless part of Q(T). This is justified once the decomposition is given. However we've also found advantages and disadvantages to each way. For the method of taking trace, it gives a minimal variance operator once the maximum of the trace is found, but we have to restrict certain operators, such as Q_X and T, to be Hilbert-Schmidt. On the other hand, the method of taking memoryless part works for a broader class of operators-operators which have decompositions, but it does not give a minimal variance operator. The best we can have is a variance operator that has a minimal memoryless part. However, there is an important aspect for taking the memoryless part. This method allows us to generalize the idea in reflexive Banach space, while the other method does not. The reason is quite simple, for it does not make sense to talk about the eigenvalue of an operator that maps from Banach space to its dual, not to mention the trace of such an operator, while it does not make sense to take the memoryless part given the decomposition. Readers are referred to Ref. 5 for the details of Wiener-Hopf filtering formulated in Hilbert resolution space. When all the problems mentioned above are solved, we would come up with the optimal filter $T_0 = C\,K^{-1}$, a causal system.

CONCLUSION

Wiener-Hopf filtering has been formulated and solved in Hilbert resolution space[5]. In this paper, we outlined the formulation in reflexive Banach resolution space and the possible way of solving it. Generalization would be accomplished once the theory of operator decomposition in Banach resolution space is completed.

REFERENCES

1. Cooper, G.R. and McGillem, C.D., Probabilistic Methods of Signal and System Analysis, New York, Holf, Rinehart and Winston, Inc., 1971.
2. Parthasarathy, K.R., Probability Measures on Metric Spaces, Academic Press, 1976.
3. Saeks, R., Resolution Space, Operators and Systems, New York, Springer-Verlag, 1973.
4. Balakrishnan, V-V., Introduction to Optimization Theory in a Hilbert Space, New York, Springer-Verlag, 1971.
5. Tung, L.J., "Random Variables, Wiener-Hopf filtering and Control Formulated in Abstract Spaces", Ph.D. Thesis, Texas Tech University, to appear.

6. Masani, P., "An Explicit Treatment of Dialation Theory", University of Pittsburgh, published notes.

7. DeSantis, R.M., "Causality Structure of Engineering Systems", Ph.D. Thesis, University of Michigan, Sept. 1971.

REPRODUCING KERNEL RESOLUTION SPACE

AND ITS APPLICATIONS II*

L. Tung and R. Saeks

*This research supported in part by AFOSR grant 74-2631.

Reproducing Kernel Resolution Space and Its Applications II*

L. Tung[+] and R. Saeks[o]

Abstract

This paper extends the concept of a reproducing kernel resolution space to a Banach space setting. The resultant reproducing kernel resolution space, however, remains a Hilbert space thereby permitting a number of problems in mathematical system theory to be transformed from Banach space to Hilbert space. Particular emphasis is given to the study of Banach space valued random variables and the scattering operator formalism for an electric network.

I. Introduction

In a previous paper[12] one of the authors exhibited the relationship between the factorization problems [19] which arise in mathematical system theory and the reproducing kernel resolution spaces introduced by Kailath and Duttweiller[13]. The purpose of the present paper is to show that much of this work can be extended to a Banach space setting without the loss of its Hilbert space character. Indeed, it is shown that the reproducing kernel resolution space for a positive self-adjoint operator mapping a reflexive Banach space to its dual is a Hilbert space. Since this is precisely the class of operators encountered when the factorization problems of mathematical system theory are extended to a Banach space setting, the resultant theory allows one to transform systems problems from a Banach space to a Hilbert space setting. In particular, it is shown that the study of certain Banach space valued random variables can be carried out in the reproducing kernal Hilbert resolution space defined by its covariance. Secondly, it is shown that an

+Dept. of EE, Univ. of Texas at El Paso, El Paso, Texas 79968.
oDept. of EE and Mathematics, Texas Tech Univ., Lubbock, Texas 79409.

*This research supported in part by AFOSR Grant 74-2631.

electric network with voltage and current vectors defined in Banach space may be characterized by a scattering operator defined on an appropriate reproducing kernel Hilbert resolution space.

A Banach resolution space is defined and its elementary properties developed. We note that the axioms for a Banach resolution space are weaker than those required for a Hilbert resolution space and, as such, the theory developed does not necessarily specialize to the classical Hilbert resolution space theory. The axioms are, however, sufficient for the present purposes. In particular, such concepts as causal, anticausal, miniphase, and maxiphase operators are well defined.

In the third section a factorization theorem for operators mapping a reflexive Banach space to its dual recently developed by Chobanian[6,7], Vakhania[3], and Masani[9] is applied to Banach resolution spaces to develop miniphase and maxiphase factorization theorems. With the help of these theorems a "unique" factor space - the generalization of the RKRS to Banach resolution space - is formulated.

While sections II and III deal with fundamental theorems, sections IV and V are devoted to the application of these theorems. The first application considered is the study of reflexive Banach space valued random variables. With the probability measure over the Banach space assumed implicitly stochastic concepts such as mean and covariance operator are defined. Since the resultant variance operator is a positive self-adjoint mapping from a reflexive Banach space to its dual the factorization theory developed in the previous section may be invoked to transform the given random variable into its reproducing kernel Hilbert resolution space.

In section V similar factorization techniques are used to define the scattering variables for an electric network characterized by voltage and current vectors taking values in a reflexive Banach space and its dual, respectively. Here, the normalizing operators take the form of maps from the given Banach space to an appropriate reproducing kernel Hilbert resolution space resulting in a scattering operator which is defined on a Hilbert space.

II. Banach Resolution Space

By a Banach resolution space, we mean a 2-tuple, $(B, {}_BF)$, where B is a Banach space and ${}_BF$ is the so-called resolution of identity in B, which is defined as follows.

Resolution of Identity in Banach Space

Def. II.1. Let B be a Banach space. By a resolution of identity, ${}_BF$, in B, we mean a family of bounded linear operators, ${}_BF(\Delta)$, on B defined for each Borel subset, Δ, of the real number set R, satisfying the following conditions:

i. ${}_BF(R) = I_B$ = identity operator on B.

ii. ${}_BF(\Delta_1)\ {}_BF(\Delta_2) = {}_BF(\Delta_1 \cap \Delta_2)$, for all $\Delta_1, \Delta_2 \in \mathcal{B}(R)$ = the set of all Borel subsets of R.

iii. ${}_BF(\bigcup_1^n \Delta_i) = \sum_1^n {}_BF(\Delta_i), \{\Delta_i\}_1^n$: finite set of disjoint Borel subsets of R.

iv. $||{}_BF(\Delta)x|| \leq ||x||$, for all $\Delta \in \mathcal{B}(R)$ and $x \in B$ (equivalently the norm of ${}_BF(\Delta)$ is either 0 or 1).

The subscript on the left in the notation, ${}_BF$, is to signify that the resolution of identity is defined in space B, and will be dropped if no ambiguity would result.

Working with a Banach resolution space $(B, {}_BF)$, it is natural and important to ask whether we can define a resoluton of identity in B^*, the dual space of B. The following theorem gives us the answer.

Thm. II.1. Let $(B, {}_BF)$ be a Banach resolution space. Then $\{{}_BF(\Delta)^* | \Delta \in \mathcal{B}(R)\}$ is a resolution of identity in B^*. This resolution of identity in B^* is called the induced resolution of identity.

The proof of this theorem follows from straightforward manipulation of the definitions of adjoint operator and resolution of identity.

The concept of a resolution of identity can be best understood by examples. Two typical examples are illustrated as follows.

Example 1. L_p: the Banach space of equivalence classes of functions that map R to R and satisfy the following inequality,

$$\int_{-\infty}^{\infty} |f(t)|^p \, dt < \infty, \text{ where } 1 < p < \infty.$$

Define

$$F(\Delta)f(t) = \chi(\Delta)f(t) = \begin{cases} 0 & , \text{ for } t \notin \Delta \\ f(t), & \text{for } t \in \Delta \end{cases}, \text{ for } f \in L_p.$$

It is easy to show that $\{F(\Delta) | \Delta \in \mathcal{B}(R)\}$ is a resolution of identity in L_p once the properties of $\chi(\Delta)$ are explored.

Example 2. Let p, q $\in$ R such that $1/p + 1/q = 1$. Then L_p is a reflexive Banach space whose dual space is L_q. For all $f \in L_p$, $g \in L_q$, define

$$(f,g) = \int_{-\infty}^{\infty} f(t)g(t) \, dt.$$

Let $\{F(\Delta) | \Delta \in \mathcal{B}(R)\}$ be the resolution of identity in L_p as defined in Example 1. Then we have

$$\begin{aligned} (F(\Delta)f, g) &= \int_{-\infty}^{\infty} [\chi(\Delta)f(t)]g(t) \, dt \\ &= \int_{\Delta} f(t)g(t) \, dt \\ &= \int_{-\infty}^{\infty} f(t)[\chi(\Delta)g(t)] \, dt \\ &= (f, \chi(\Delta)g) \end{aligned}$$

So $\{\chi(\Delta) | \Delta \in \mathcal{B}(R)\}$ is also the induced resolution of identity in L_q.

Causality of Operators

Def. II.2. Let $(X, {}_XF)$, $(Y, {}_YF)$ be Banach resolution spaces. $T : X \to Y$, is a linear bound operator.

i. T is said to be causal, if

$${}_XF^t x_1 = {}_XF^t x_2 \Rightarrow {}_YF^t T x_1 = {}_YF^t T x_2, \text{ where}$$

$${}_BF^t = {}_BF(-\infty, t), \; B = X, Y.$$

ii. T is said to be anti-causal if

$$ {}_XF_t\, x_1 = {}_XF_t\, x_2 \Rightarrow {}_YF_t\, T\, x_1 = {}_YF_t\, T\, x_2, \text{ where } {}_BF_t = I_B - {}_BF^t,\ B = X, Y. $$

iii. T is said to be memoryless if T is causal and anti-causal.

iv. T is said to be left-miniphase if

$$ {}_YF^t\, T\, x = 0 \Leftrightarrow {}_XF^t\, x = 0, $$

v. T is said to be left-maxiphase if

$$ {}_YF_t\, T\, x = 0 \Leftrightarrow {}_XF_t\, x = 0, $$

vi. T is said to be right-miniphase if

$$ \overline{T[{}_XF_t[X]]} = {}_YF_t[Y], $$

vii. T is said to be right-maxiphase if

$$ \overline{T[{}_XF^t[X]]} = {}_YF^t[Y]. $$

Based on the above definitions, the following results may be readily verified?

1. If T is miniphase, left- or right-, then T is causal.

2. If T is maxiphase, left- or right-, then T is anti-causal.

3. When the spaces involved are reflexive and $T\ X \rightarrow Y$ is bounded and linear, then

   (a) T is causal iff T* is anti-causal
   (the induced resolutions of identity have been used for the determination of the causality of T*)

   (b) T is left-miniphase iff T* is right-maxiphase

4. When the spaces involved are reflexive and $T: X \rightarrow Y$ is invertible.

   (a) T is left-miniphase(maxiphase) if it is right-miniphase (maxiphase).

   (b) T is miniphase $\Rightarrow$ T & T^{-1} are causal
   T is maxiphase $\Rightarrow$ T & T^{-1} are anti-causal.

The reader is referred to Ref. 2 for the details.

III. Operator Factorization and RKRS

A number of applications arise in system theory wherein it is desired to factor an operator, Q, either in the form KK* or T*T. For an operator to be factorized in either of these forms it has to be "positive" and "self-adjoint". Although undefined in general, these two commonly used properties for operators over Hilbert space can be extended to operators that map a reflexive Banach space to its dual space. They are defined as follows:

Def. III.1. i. B is a reflexive Banach space.

ii. $T: B \to B^*$, is linear and bounded.

T is said to be positive if

$$(x, T x) \geq 0, \text{ for all } x \in B.$$

T is said to be self-adjoint if

$$T^* = T.$$

Note that $T^*: B^{**}=B \to B^*$, so that it makes sense to equate T and T^*.

For positive and self-adjoint operators, we have the following theorem. The theorem is stated without proof. Interested readers are referred to Masani's work (9).

Thm. III.1. Let $Q: B \to B^*$, be a linear, bounded, positive and self-adjoint operator, where B is a reflexive Banach space and B^* is its dual. Then there exists a Hilbert space H and a linear bounded operator K mapping H to B^* such that $Q = K K^*$.[6,7,8,9]

Left- and Right- Factorization

Although Thm. III.1 yield a factorization through a Hilbert space, which is valid for any linear, bounded, positive, self adjoint operator mapping a reflexive Banach space to it dual our applications in System Theory require that the factors, K, have appropriate causality properties. Based on Thm.III.1, we construct appropriate resolutions of identity in the spaces involved which yield the following theorems.

Thm.III.2(left-factorization).

i. (B, F) is a reflexive Banach resolution space.

ii. $Q: (B, F) \to (B^*, F^*)$, where F^* is the induced resolution of identity in B^*. Q is linear, bounded, positive and self-adjoint.

Then there exists a Hilbert resolution space $(H,E)^\dagger$ and a linear bounded operator $K: (H, E) \to (B^*, F^*)$, such that

1. $Q = K K^*$,
2. K is left-miniphase,
3. The factorization is unique up to a memoryless unitary transformation.

† The requirements for a Hilbert resolution space are more restrictive than those for a Banach resolution space, namely $E(\Delta)^* = E(\Delta)$ is required for a Hilbert resolution space, which condition does not make sense in a Banach resolution space.

Thm.III.3(right-factorization).

i. $Q: (B, F) \to (B^*, F^*)$.

ii. Q is linear, bounded, positive and self-adjoint. Then there exists a Hilbert resolution space $(\underset{\sim}{H}, \underset{\sim}{E})$ and a linear bounded operator T: $(B, F) \to (\underset{\sim}{H}, \underset{\sim}{E})$, such that

1. $Q = T^* T$,
2. T is right-miniphase,
3. The factorization is unique up to a memoryless unitary transformation.

The proofs for the above theorems are time-consuming and they are not our prime concern. Hence, they will not be presented here. Interested readers are referred to the Appendix for their proofs.

Reproducing Kernel Resolution Space (RKRS)

There is a common statement in each theorem of the previous paragraph, i.e. "The factorization is unique up to a memoryless unitary transformation". For certain applications such as the study of Banach space valued random variable, we would like to eliminate this ambiguity. This is achieved via the concept of a reproducing kernel resolution space. First, we define some notation.

Def.III.2. Q, K, ant T are defined as in the previous paragraph i.e. $Q = K K^* = T^* T$. Let $\tilde{H}_Q = R(K)^+$, $\underset{\sim}{H}_Q = R(T^*)^+$. For x, y in R(K), let

$$(x, y)_{\tilde{H}_Q} = (K^{-L} x, K^{-L} y)_{\tilde{H}}, \text{ and for w, z in } R(T^*), \text{ let}$$

$$(z, w)_{\tilde{H}_Q} = (T^{*-L} z, T^{*-L} w)_{\tilde{H}}.$$

Define ${}_Q\tilde{E}^t = K \tilde{E}^t K^{-L}$ on the $\tilde{H}_Q$ and $Q^{E}_{\sim}t = T^* \underset{\sim}{E}_t T^{*-L}$ on $\underset{\sim}{H}_Q$, then extend them to B(R). Note that E corresponds to E in Thm.III.2.

+ R(X) denotes the range of the operator, X.

Then, we have the following theorem.

Thm.III.4. $(\tilde{H}_Q, {}_Q\tilde{E})$ (the so-called RKRS) defined above is a Hilbert resolution space which is independent of the factorization $Q = K K^*$ used in its definition. Moreover,

i. $(\tilde{H}_Q, {}_Q\tilde{E})$ is unitarily equivalent to $(\tilde{H}, \tilde{E})$,

ii. $R(Q)$ is dense in $\tilde{H}_Q$,

iii. $(F^t)^* x = 0$ iff $\tilde{E}^t x = 0$, for $x \in \tilde{H}_Q$,

iv. $K: (\tilde{H}, \tilde{E}) \to (\tilde{H}_Q, {}_Q\tilde{E})$, is memoryless,

where $(\tilde{H}, \tilde{E})$ corresponds to (H, E) in Thm. III. 2.

Proof: i. $\tilde{H}_Q = R(K)$, so $\tilde{H}_Q$ is a linear vector space.

ii. By definition, $(x, y)_{\tilde{H}_Q} = (K^{-L} x, K^{-L} y)_{\tilde{H}}$. It is trivial to show that

1. $(x, y) = (y, x)$, for all $x, y \in \tilde{H}_Q$.
2. $(x, ay) = a(x, y)$, for all $a \in R$, $x, y \in \tilde{H}_R$
3. $(x, y + z) = (x, y) + (x, z)$ for all $x, y, z \in \tilde{H}_Q$,
4. $(x, x)_{\tilde{H}_Q} = (K^{-L} x, K^{-L} x)_{\tilde{H}} \geq 0$, for all $x \in \tilde{H}_Q$ and since K is linear, $x \neq 0$ implies $K^{-L} x \neq 0$. So $(x, x) > 0$, for $x \in \tilde{H}_Q$ and $x \neq 0$.

So $(x, y)_{\tilde{H}_Q}$ is an inner product over $\tilde{H}_Q$.

iii. Let $\{x_i\}$ be a Cauchy sequence in $\tilde{H}_Q$, then $\{K^{-L} x_i\}$ is Cauchy in $\tilde{H}$. So $K^{-L} x_i \to z$, for some $z \in \tilde{H}$. But $z = K^{-L} K z$, so $x_i \to K z$ in $\tilde{H}$. Hence $\tilde{H}_Q$ is a Hilbert space.

iv. Assume $Q = K K^* = K' K'^*$, both K and K' are left-miniphase factorizations of Q on factor spaces $(\tilde{H}, \tilde{E})$ and $(\tilde{H}', \tilde{E}')$, respectively.

1. By Thm.III.2, there exists a memoryless unitary transfor-

mation $U: (\tilde{H}, \tilde{E}) \rightarrow (\tilde{H}', \tilde{E}')$, such that $K\,U = K'$. So $R(K') = R(K\,U) = R(K)$, since U is onto. Hence $\tilde{H}_Q$ is independent of the factorization.

2. $$(K'^{-L} z, K'^{-L} w)_{\tilde{H}'} = ((K\,U)^{-L} z, (K\,U)^{-L} w)_{\tilde{H}'}$$
$$= (U^{-1} K^{-L} z, U^{-1} K^{-L} w)_{\tilde{H}'}$$
$$= (K^{-L} z, K^{-L} w)_{\tilde{H}}, \text{ for all}$$
$z, w \in \tilde{H}_Q$. So the inner product is independent of the factorization.

3. $$K' \tilde{E}'^{t} K'^{-L} = (K\,U) \tilde{E}'^{t} (K\,U)^{-L} = K\,U\,\tilde{E}'^{t} U^{-1} K^{-L}$$
$$= K\,E^{t}\,U\,U^{-1} K^{-L} = K\,E^{t} K^{-1}$$

Therefore, ${}_QE$ is also independent of the factorization.

v. Define $\tilde{K}: \tilde{H} \rightarrow \tilde{H}_Q$ by $\tilde{K}\,x = K\,x$, for all $x \in \tilde{H}$. Then $\tilde{K}$ is 1-1 and onto and for all $x \in \tilde{H}$,

$$||\tilde{K}\,x||^2_{\tilde{H}_Q} = ||K\,x||^2_{\tilde{H}_Q} = ||K^{-L} K\,x||^2_{\tilde{H}} = ||x||^2_{\tilde{H}}$$

So $\tilde{K}$ is a unitary mapping. Furthermore, we have ${}_Q\tilde{E}^{t} = K\,\tilde{E}^{t} K^{-L}$ on $\tilde{H}_Q$ and for all $z \in \tilde{H}_Q$, there is a unique $x \in \tilde{H}$ such that $z = K\,x = \tilde{K}\,x$. So $x = K^{-L} z = \tilde{K}^{-1} z$, i.e. $K^{-L} = \tilde{K}^{-1}$. Hence ${}_Q\tilde{E}^{t} = \tilde{K}\,\tilde{E}^{t}\,\tilde{K}^{-1}$. This means that $(\tilde{H}_Q, {}_Q\tilde{E})$ is unitarily equivalent to $(\tilde{H}, \tilde{E})$.

vi. $R(Q) = R(K\,K^*) = K[R(K^*)] = \tilde{K}[R(K^*)]$
Since $R(K^*)$ is dense in $\tilde{H}$, so $\tilde{K}[R(K^*)]$ is dense in $\tilde{H}_Q$ (for $\tilde{K}$ is unitary).

vii. 1. K is left-miniphase, so
$$\tilde{E}^{t}\,x = 0 \text{ iff } (F^{t})^* K\,x = 0, \text{ for } x \in \tilde{H}.$$

2. For all $z \in \tilde{H}_Q$, there exists $x \in \tilde{H}$ such that $z = K\,x$ or $x = K^{-L} z$. So

$\tilde{E}^t K^{-L} z = 0$ iff $(F^t)^* z = 0$, $z \in \tilde{H}_Q$.

3. If $(F^t)^* z = 0$, then

$_Q\tilde{E}^t z = K \tilde{E}^t K^{-L} z = 0.$

And if $_Q\tilde{E}^t z = 0$, i.e. $K \tilde{E}^t K^{-L} z = 0$, then

$\tilde{E}^t K^{-L} z = 0$. So $(F^t)^* z = 0$.

viii. K: $(\tilde{H}, \tilde{E}) \to (\tilde{H}_Q, {}_Q\tilde{E})$, is actually $\tilde{K}$ defined in v. K is left-miniphase, so $\tilde{E}^t x = 0$ iff $(F^t)^* K x = 0$. Hence $\tilde{E}^t x = 0$ iff $_Q\tilde{E}^t K x = 0$, for $x \in H$. But since $\tilde{K} = K$, the above equation implies $\tilde{K}$ is also a left-miniphase. So $\tilde{K}^* = \tilde{K}^{-1}$ is causal (by being left-miniphase), and $\tilde{K}$ is also anti-causal. Therefore, $\tilde{K}$ is memoryless. #

Following immediately from the previous theorem is a rather interesting result which gives us a sort of "unique" left-factorization. This result is indicated in the next corollary.

Cor. Let $(\tilde{H}_Q, {}_Q\tilde{E})$ be defined as in Thm.III.4. Then there exists a left-factorization P: $(\tilde{H}_Q, {}_Q\tilde{E}) \to (B^*, F^*)$, such that

i. $P z = z$ for all $z \in \tilde{H}_Q \subseteq B^*$,

ii. $P^* b = Q b$, for all $b \in B$.

Proof: By Thm.III.4, we have $\tilde{K}: \tilde{H} \to \tilde{H}_Q$, a memoryless unitary operator, and $\tilde{K} x = K x$, for all $x \in \tilde{H}$. Define

$P = K \tilde{K}^{-1}: \tilde{H}_Q \to B^*$. Then

1. For all $z \in \tilde{H}_Q$, there exists $x \in \tilde{H}$ such that $\tilde{K} x = z$.

2. By Thm.III.4,

$_Q\tilde{E}^t z = 0$ iff $(F^t)^* z = 0$, for $z \in \tilde{H}_Q$. So

$_Q\tilde{E}^t z = 0$ iff $(F^t)^* P z = 0$, for $z \in \tilde{H}_Q$. Hence P is

left-miniphase.

3. $P^* b = (K \tilde{K}^{-1})^* b = (\tilde{K}^{-1})^* K^* b = \tilde{K} K^* b = K K^* b = Q b$,

for all $b \in B$. And $P P^* b = P (Q b) = Q b$. #

The "uniqueness" we mentioned is due to the fact that $(\tilde{H}_Q, {}_Q\tilde{E})$ is independent of the factorization.

As in the previous paragraph, there are corresponding dual theorems to Thm. III.4. These theorems are described below, with proofs only sketched.

Thm.III.5. $(\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$, defined at the beginning of this section, is a Hilbert space which is independent of the factorization, $Q = T^* T$, used in its definition. Moreover,

i. $(\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$ is unitarily equivalent to $(\underset{\sim}{H}, \underset{\sim}{E})$,

ii. $R(Q)$ is dense in $\underset{\sim}{H}_Q$,

iii. $(F^t)^* x = 0$ iff $\underset{\sim}{E}^t x = 0$, for $x \in \underset{\sim}{H}$,

iv. $T^*: (\underset{\sim}{H}, \underset{\sim}{E}) \to (\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$, is memoryless.

Proof: i. The proof of $(\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$ being a Hilbert resolution space and being independent of the factorization is essentially the same as that of Thm.III.4, and therefore is omitted.

ii. Define $\tilde{T}^* : (\underset{\sim}{H}, \underset{\sim}{E}) \to (\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$, by $\tilde{T}^* x = T^* x$ for all $x \in \underset{\sim}{H}$. Then $\tilde{T}^*$ is 1-1 and onto. For all $x \in \underset{\sim}{H}$,

$$||\tilde{T}^* x||^2_{\underset{\sim}{H}_Q} = ||T^* x||^2_{\underset{\sim}{H}_Q} = ||(T^*)^{-L} T^* x||^2_{\underset{\sim}{H}} = ||x||^2_{\underset{\sim}{H}}.$$

So $\tilde{T}^*$ is a unitary mapping. With $\tilde{T}^*$ defined as above, it is routine to verify the rest of the theorem. But we need to note that $T^*: (\underset{\sim}{H}, \underset{\sim}{E}) \to (\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$ in iv., is actually the $\tilde{T}^*$ defined above.

Cor. ($\underset{\sim}{H}_Q$, ${}_Q\underset{\sim}{E}$) defined as above, then there is a right-factorization (mini-phase) of Q, q: (B, F) → ($\underset{\sim}{H}_Q$, ${}_Q\underset{\sim}{E}$), such that

i. $q\ x = Q\ x$, for all $x \in B$,

ii. $q^*\ z = z$, for all $z \in \underset{\sim}{H}_Q \subseteq B^*$.

Proof: By Thm.III.5, $\tilde{T}^*: \underset{\sim}{H} \to \underset{\sim}{H}_Q$, is memoryless and unitary.

Define $q = \tilde{T}^*\ T: (B, F) \to (\underset{\sim}{H}_Q, {}_Q\underset{\sim}{E})$. Then

i. $q\ x = \tilde{T}^*\ T\ x = T^*\ (T\ x) = Q\ x$, for all $x \in B$.

ii. $q^*\ z = (\tilde{T}^*\ T)^*\ z = T^*\ \tilde{T}\ z = \tilde{T}^*(\tilde{T}\ z) = z$, for all $z \in \underset{\sim}{H}_Q$, since $\tilde{T}^*$ is unitary.

iii. By Thm.III.5,

${}_Q\underset{\sim}{E}_t\ x = 0$ iff $(F_t)^*\ x = 0$, for $x \in \underset{\sim}{H}$. So

${}_Q\underset{\sim}{E}_t\ x = 0$ iff $(F_t)^*\ q^*\ x = 0$, for $x \in \underset{\sim}{H}$. So q^* is left-maxiphase, i.e. q is right-miniphase.

iv. $q^*\ q\ b = (\tilde{T}^*\ T)^*\ (\tilde{T}^*\ T)\ b = T^*\ \tilde{T}\ \tilde{T}^*\ T\ b = T^*\ T\ b = Q\ b$, for all $b \in B$.

IV. Banach Space Valued Random Variables

One way to view a random process is to consider it as a random variable which takes values in a function space. Of course, we have to use an adequate probability measure to make the idea work. Fortunately, this kind of measure has been defined for metric spaces.[3,4] In this section, we first define stochastic properties such as "mean" and "variance operator" for a Banach space valued random variable. We then look into the factorization of the variance operator and the results that can be derived therefrom, i.e. the RKRS. Interestingly enough, the RKRS of a Banach space valued random variable is a Hilbert space. This seems to be a nice result, but there are obstacles for further application. All these will be discussed in the following.

(1) Covariance Operator

A probability measure on Banach space is a rather complicated matter. In the sequel, we implicitly assume its existence as indicated by the expected value symbol $E\{\cdot\}$. The reader is referred to reference 4 for the details.

Let ρ, π denote finitely additive random variables taking values in a relexive Banach space B. Assume

$$\begin{array}{ll} \text{i.} & E\{|(\rho, x^*)|\} < \infty, \text{ for all } x^* \in B^* \\ \text{ii.} & E\{(\rho, x^*)\} \text{ is continuous in } x^*. \end{array} \qquad \text{(a)}$$

Then there exists a unique $m_\rho \in B$ such that

$$E\{(\rho, x^*)\} = (m_\rho, x^*)$$

Since $E\{(\rho,x^*)\}$ is a continuous linear functional on B^*, so it is represented by an element of $B^{**} = B$. m_ρ is termed the mean of the random variable ρ. It has the following properties

i. $m_{\rho+\pi} = m_\rho + m_\pi$,

ii. $||m_\rho|| \leq E\{||\rho||\}$,

iii. If $L: B \to B$, is bounded and linear, then

$$m_{L\rho} = L\, m_\rho.$$

As in most stochastic processes, the mean is not our prime concern. In the sequel we thus assume that all random variables have zero mean. For the definition of the variance operator, we have to assume the following.

$$\begin{array}{ll} \text{i.} & E\{|(\rho, x^*)\,(\pi, y^*)|\} < \infty, \text{ for all } x^*, y^* \in B^*, \\ \text{ii.} & E\{(\rho, x^*)\,(\pi, y^*)\} \text{ is continuous in } x^* \text{ and } y^*. \end{array} \qquad \text{(b)}$$

It is easy to show that condition (b) implies condition (a). Furthermore, we have the following lemma to facilitate the definition of the variance operator.

Lemma. A continuous bilinear functional, $(x|y)$, on a Banach space B is also bounded (i.e. there exists $M \in R$ such that

$|(x|y)|/(||x||\cdot||y||) < M$, for all $x, y \in B$).

Now if we fix y^*, then $E\{(\rho, x^*)\ (\pi, y^*)\}$ is a bounded linear functional on B^* (hence an element of $B^{**}=B$). This implies that there exists a unique $p_{y^*} \in B$ such that

$E\{(\rho, x^*)\ (\pi, y^*)\} = (p_{y^*}, x^*)$ for all $x^* \in B^*$.

If we now define a mapping $Q_{\rho\pi}: B^* \to B$ by

$$Q_{\rho\pi}\, y^* = p_{y^*}.$$

it can easily be verified that $Q_{\rho\pi}$ is linear. Morever, $Q_{\rho\pi}$ is bounded, since

$$||Q_{\rho\pi}\, y^*|| = ||p_{y^*}|| = \sup_{x^*} \frac{|E\{(\rho, x^*)\ (\pi, y^*)\}|}{||x^*||}$$

$$\leq \sup_{x^*} \frac{M||x^*||\cdot||y^*||}{||x^*||} = M||y^*||.$$

The operator $Q_{\rho\pi}$ is termed the covariance operator of the random variables ρ and π.

Covariance operators satisfy the following conditons:

i. $Q_{(L\rho)\,(K\pi)} = L\, Q_{\rho\pi}\, K^*$, where L and K are bounded and linear operators on B.

ii. Let $Q_\rho = Q_{\rho\rho}$; then

$$Q_{\rho+\pi} = Q_\rho + Q_{\rho\pi} + Q_{\pi\rho} + Q_\pi.$$

Q_ρ is called the variance operator of ρ.

iii. $Q_{\rho\pi} = Q_{\pi\rho}^*$, in particular, $Q_\rho^* = Q_\rho$.

iv. Q_ρ is positive; i.e. $(Q_\rho\, y^*, y^*) = E\{(\rho, y^*)^2\} \geq 0$.

(2) <u>RKRS For a Banach Space Valued Random Variables</u>

As mentioned in the previous paragraph, a reflexive Banach space valued random variable has a variance operator Q_ρ which is positive and self-adjoint. Then Thm.III.2

and Thm.III.4 come into the picture and we have the following: There exists a Hilbert resolution space (H,E) and a left-factorization (miniphase) $K: (H,E) \to (B,F)$ such that $Q_\rho = K K^*$. Moreover, the RKRS, $(H_\rho, {}_\rho E)$, which corresponds to $(\tilde{H}_Q, {}_Q\tilde{E})$ in Thm.III.4 is also assured to exist. Since we have $R(Q_\rho) \subseteq H_\rho \subseteq B$, one of the natural questions to ask is whether the random variable takes values *only* in H_ρ, and, if it does, what can we say about the original random variable. The answer to the first part of the question is no, and a counter-example has been constructed.[2] However, if we happen to have the random variable taking values in H_ρ, we would have the following properties. First define $B_0 = \overline{R(K)}^B$, where $K: H \to B$. Then $K^{-L}: B_0 \to H$ and we have $(K^{-L})^*: H \to B_0^*$. Consider $E\{(\rho, x)_\rho (\rho, y)_\rho\}$, for x, y in H_ρ.

$$E\{(\rho, x)_\rho (\rho, y)_\rho\} = E\{(K^{-L} \rho, K^{-L} x)_H (K^{-L} \rho, K^{-L} y)_H\}$$

$$= E\{(\rho, (K^{-L})^* K^{-L} x)B_0 \ (\rho, (K^{-L})^* K^{-L} y)B_0\}.$$

Note: $(K^{-L})^* K^{-L} x$ and $(K^{-L})^* K^{-L} y$ are elements of B_0^*, i.e. linear functionals on B_0. By the Hahn-Banach theorem[5,11], there exist x^* and y^* in B^* such that

$$x^*|_{B_0} = (K^{-L})^* K^{-L} x,$$

and

$$y^*|_{B_0} = (K^{-L})^* K^{-L} y.$$

So

$$(\rho, (K^{-L})^* K^{-L} x)_{B_0} = (\rho, x^*)_B$$

and

$$(\rho, (K^{-L})^* K^{-L} y)_{B_0} = (\rho, y^*)_B.$$

Therefore,

$$E\{(\rho, x)_\rho (\rho, y)_\rho\} = E\{(\rho, x^*)_B (\rho, y^*)_B\}$$

$$= (Q_\rho x^*, y^*)_B$$

$$= (KK^* x^*, y^*)_B = (K^* x^*, K^* y^*)_H$$

$$= (K^{-6} x, K^{-6} y)_H = (x, y)_\rho$$

as such the random variable

ρ has the identity operator on H as its variance operator. The whole idea can be explained by the following example:

Example 3. Let $r = 5/12$ and $h = (1, (1/2)^r, (1/3)^r, \ldots) \in \ell_3$, where ℓ_3 = {real number sequence $\{y_i\} \mid \sum_{i=1}^{\infty} y_i^3$ is finite}. Let x be a zero-mean random variable taking values in the real number set R and let its variance be 1. The $\rho = xh$ is a random variable taking values in ℓ_3. The dual space of ℓ_3 is $\ell_{3/2}$ and ℓ_3 is reflexive.

(i) For all $z \in \ell_{3/2}$,

$$E\{(\rho,z)\} = E\{(xh, z)\} = E\{x(h,z)\}$$
$$= (h, z)\, E\{x\} = 0.$$

Hence ρ is zero-mean.

(ii) For all $z, w \in \ell_{3/2}$,

$$E\{(\rho, z)(\rho,w)\} = E\{(xh,z)(xh,w)\} = E\{x^2(h,z)(h,w)\}$$
$$= (h,z)(h,w)\, E\{x^2\} = (h,z)(h,w)$$
$$\overset{\text{def}}{=} (Q_\rho z, w).$$

Thus $Q_\rho z = (h,z)h$.

(iii) Since the range of Q_ρ is 1-dimensional, the possible factor space is the real set R. Let $K^* = \ell_{3/2} \to R$ be the functional $(h,\cdot)$.
then $K = R \to \ell^*_{3/2} = \ell_3$ is defined by the following equation,

$(Kd,z) = (d, K^* z)$, for all $d \in R$ and $z \in \ell_{3/2}$.

Since

$$(d, K^* z) = d\, K^* z = d(h,z) = (dh,z),$$

we have

$$Kd = dh, \text{ for all } d \in R.$$

For all $z \in \ell_{3/2}$, we have

$$KK^* z = K(h,z) = (h,z)\, h = Q_\rho z.$$

Hence, $Q_\rho = KK^*$

(iv) Furthermore, we have

$$R^t \underline{\text{def}} \ \overline{R(K^*F^t)} = \begin{cases} R & , \text{for } t > 0 \\ \{0\} & , \text{for } t < 1 , \end{cases}$$

for

$$R(K^*F^t) = K^*F^t [\ell_{3/2}]$$
$$= \{K^*F^t z \mid z \in \ell_{3/2}\}$$
$$\{(h, F^t z) \mid z \in \ell_{3/2}\}$$
$$\begin{cases} R & , \text{for } t > 1 \\ \{0\} & , \text{for } t < 1. \end{cases}$$

Hence, E^t on R is the step function

$$U(t-1) = \begin{cases} 1 & , t > 1 \\ 0 & , t < 1 \end{cases}$$

(v) a. When $E^t d = 0$,

case 1 $(t < 1) = F^t = 0$, so $F^t Kd = 0$;

case 2 $(t > 1) = E^t = I_R = 1$, so $d = 0$.

Thus $F^t K d = 0$.

b. When $F^t K d = 0$,

case 1 $(t < 1) = E^t = 0$, so $E^t d = 0$;

case 2 $(t > 1) = F^t K d = F^t dh = d F^t h = 0$

However, $F^t h = (1, (1/2)^r, (1/3)^r, \dots (1/i)^r, 0, \dots)$

$\neq 0$, where $1 \leq i \leq t$.

Hence, $d = 0$. Thus, $E^t d = 0$.

From a and b, K is a left-miniphase.

(vi) $H_Q = K[R] = \{dh \mid d \in R\}$

For $d_1 h, d_2 h \in H_Q$, the inner product is defined as

$$(d_1 h, d_2 h)_Q = d_1 d_2$$

(vii) Since ρ takes values in H_Q only, ρ can be considered as a zero-mean random variable over H_Q with the identity mapping on H_Q as its variance. This statement can be verified easily.

Note, since the stochastic character of the given random variable, ρ, was derived from the scalar random variable, x, it is appropriate that ρ can be characterized completely in terms of its representation in the one dimensional RKRS, H_ρ.

V. Scattering Operator

Classically in network analysis, network variables, such as voltage and current, are assumed to be in Hilbert space. Although the scattering variables are a very useful tool in network analysis, the significance of the normalizing impedance used in their derivation is not clear. Situations have occurred where we have to assume that the network variables are defined in Banach space. If the scattering variables are to be well defined here, the function of the normalizing impedance should be the transformation of network variables defined in Banach space into elements of a Hilbert space. Theoretically, it is much easier to work with Hilbert space. Therefore, the significance of the normalizing impedance lies in the fact that it transforms a problem defined in Banach space into a Hilbert space problem. In this section, we will extend the idea of scattering variables to networks with their voltage and current variables defined in Banach spaces with the help of the factorization theorems developed in Chapter III.

Thinking of the impedance Z, or the transfer function, as an operator from a current space to a voltage space, the power $V.I$ is a scalar quanity. Here V denotes voltage and I current, and the power equality implies that the voltage plays the role of a linear functional operating on the current. Thus, Z may naturally be viewed as a mapping from a current space to its dual, a voltage space. Similarly, an admittance assumes a dual role mapping voltage to current. In that case, "I" is a linear functional over the voltage space. What model could be better than a reflexive Banach space to fit the structure? Therefore, in this work, the current and voltage spaces are chosen to be dual reflexive Banach spaces.

Unlike the classical case, the normalizing impedance (or admittance) operator is not invertible, in general. Besides being positive, causal, linear, and bounded, we, however, also assume that the normalizing impedance is 1-1. For this class of operator, we have the following special factorization theorems.

Thm. V.1. Let $Z: (B, F) \rightarrow (B^*, F^*)$, where B is a reflexive Banach space, be positive, causal, 1-1, linear and bounded. Let $M = 1/2 \cdot (Z + Z^*)$; then M is also positive and furthermore, self-adjoint. Then

i. There exists a Hilbert resolution space $(H, \tilde{E})$ and a left-factorization of M, $K_0: (H, \tilde{E}) \rightarrow (B^*, F^*)$, such that K_0 is left-miniphase.

ii. There exists a Hilbert resolution space $(H, \underset{\sim}{E})$ and a right-factorization of M, $T_0: (B, F) \rightarrow (H, \underset{\sim}{E})$, such that T_0 is right-miniphase.

Proof: The existence of the left- and right-factorization follows from Thm. III.2 and Thm. III.3.

Note here that we use the same Hilbert space for left- and right-factorization. This can be justified from the proofs of Thm.III.2 and Thm.III.3 in Appendix.

Thm. V.2. Let M be defined as above. If $K: (H, \tilde{E}) \rightarrow (B^*, F^*)$, is a linear bounded operator such that

i. K is 1-1 and causal,

ii. $M = K K^*$,

then there exists a linear bounded operator $U: (H, \tilde{E}) \rightarrow (H, \tilde{E})$, such that

a. $K = K_0 U$, (K_0 is as defined in Thm.V.1.)

b. U is causal and unitary.

Proof: (i) For all $y \in H$ such that $Ky \in Ko[H]$,

define $Uy = Ko^{-L} K y$.

For $b \in B$, we have

$KK^*b = KoKo^*b \in Ko[H]$.

Hence U is defined over $K^*[B]$ which is dense in H.

For all $y \in K^*[B]$, there exists $x \in B$ such that $y = K^* x$ and we have

$$||Uy||^2_H = ||Ko^{-L} K y||^2_H = ||Ko^{-L} K K^* x||^2_H$$

$$= ||Ko^{-L} Ko Ko^* x||^2_H = ||Ko^* x||^2_H$$

$$= (Ko^* x, Ko^* x)_H$$

$$= (x, Ko Ko^* x)_B$$

$$= (x, K K^* x)_B$$

$$= (K^* x, K^* x)_H$$

$$= ||K^* x||^2_H = ||y||^2_H$$

Therefore, U is isometric on $K^*[B]$. U can thus be extended over H isometrically.

(ii) Similarly, define

$\overline{V} y = K^{-L} Ko y$ over $Ko^*[B]$.

$\overline{V}$ is isometric and can be extended over H isometrically.

(iii) For all $h \in H$, there exists sequence $\{xi\}$ in B such that

$Ko^* xi \to h$.

This implies

$\overline{V} (Ko^* xi) \to \overline{V} h$

and

$U \overline{V} (Ko^* xi) \to U \overline{V} h$

However,

$$\bar{V}\,(Ko^{*}\,x_i) = K^{-L}\,Ko\,Ko^{*}\,x_i$$
$$= K^{-L}\,K\,K^{*}\,x_i$$
$$= K^{*}\,x_i$$

Hence

$$U\,\bar{V}\,(Ko^{*}\,x_i) = Ko^{-L}\,K\,K^{*}\,x_i$$
$$= Ko^{-L}\,Ko\,Ko^{*}\,x_i$$
$$= Ko^{*}\,x_i \rightarrow h.$$

Therefore, $U\,\bar{V}\,h = h$ and U is an onto mapping. Hence, U is unitary.

(iv) For all $h \in H$, there exists sequence $\{x_i\}$ such that $K^{*}\,x_i \rightarrow h$.

We also have,

$$Ko\,U\,K^{*}\,x_i = Ko\,Ko^{-L}\,K\,K^{*}\,x_i = Ko\,Ko^{-L}\,Ko\,Ko^{*}\,x_i$$
$$= Ko\,Ko^{*}\,x_i = K\,K^{*}\,x_i \rightarrow K\,h,$$

but

$$Ko\,U\,K^{*}\,x_i \rightarrow Ko\,U\,h.$$

Hence

$$Ko\,U\,h = K\,h \text{ for all } h \in H, \text{ i.e. } Ko\,U = K.$$

(v) Since $Ko\,U = K$, we have

$$(F^{t})^{*}\,Ko\,E^{t}\,U = (F^{t})^{*}\,Ko\,U, \text{ (Ko is causal)}$$
$$= (F^{t})^{*}\,K$$
$$= (F^{t})^{*}\,K\,E^{t}, \text{ (K is causal)}$$
$$= (F^{t})^{*}\,Ko\,U\,E^{t}$$
$$= (F^{t})^{*}\,Ko\,E^{t}\,U\,E^{t}, \text{ (Ko is causal)}$$

Hence,

$$(F^{t})^{*}\,Ko\,E^{t}\,[E^{t}\,U - E^{t}\,U\,E^{t}] = 0.$$

This implies that $E^{t}\,U - E^{t}\,U\,E^{t} = 0$, since Ko is left-miniphase.

Therefore, U is causal.

Thm.V.3. Let M be defined as above. If $T = (B, F) \to (H, \underset{\sim}{E})$ is a linear bounded operator such that i. T is causal and has dense range.

ii. $M = T^* T$

Then there exists $W = (H, \underset{\sim}{E}) \to (H, \underset{\sim}{E})$ such that

a. $T = W\, To$

b. W is causal and unitary.

Proof: (i) Define $W^* y = (To^*)^{-L} T^* y$, for $y \in T[B]$.

By similar argument as in Thm.V.2., $\overline{W}^*$ can be proved to be isometric on T[B]. Since T[B] is dense in H, $\overline{W}^*$ can be extended to H isometrically. Also like in Thm.V.2, $\overline{W}^*$ can be proved to be unitary.

(ii) Also by similary argument, it can be proved that

$To^* \overline{W}^* = T^*$. Hence $T = \overline{W}\, To$

(iii)
$$\begin{aligned}(F_t)^* To^* E_t \overline{W}^* &= (F_t)^* To^* \overline{W}^*, \text{ (To* is anti-causal)}\\ &= (F_t)^* T^*\\ &= (F_t)^* T^* E_t, \text{ (T* is anti-causal)}\\ &= (F_t)^* to^* \overline{W}^* E_t\\ &= (F_t)^* To^* E_t \overline{W}^* E_t\end{aligned}$$

Hence $(F_t)^* To^* (E_t \overline{W}^* - E_t \overline{W}^* E_t) = 0$.

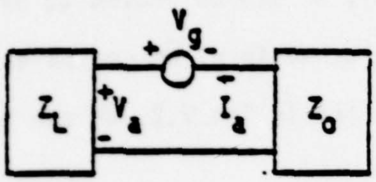
This implies that $E_t \overline{W}^* - E_t \overline{W}^* E_t = 0$, since To^* is left-maxiphase.

Therefore, $\overline{W}^*$ is anti-causal, i.e.

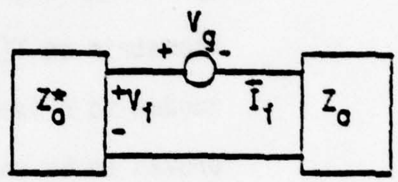
$\overline{W}$ is causal.

It is trivial to show that if we have a causal and unitary operator U on $(H, \tilde{E})$ then Ko U is causal and $(KU)(KU)^* = M$. Similarly, if we have a causal and unitary operator $\overline{W}$ on $(H, \mathcal{E})$, then $\overline{W}$ To is causal and $(\overline{W}\ To)^*(\overline{W}\ To) = M$.

The significance of these facts is that they enable us to choose causal and unitary operators U & W as desired in order to make the factorization satisfy some additional requirements. Unfortunately, the proof of existence and the construction of these U & W is currently beyond our reach, even though it is trivial to do so in the classical case.[10] With this in mind, let us now consider the following network:



a. n-port network series loaded



b. optimal matching situation

Figure 1.

In the figure, we have

1. $I_a, I_i \in B$ = a reflexive Banach space
2. $V_a, V_g, V_i \in B^*$
3. $Z_o, Z_L = (B, F) \to (B^*, F^*)$, linear and bounded.

Although the circuit diagram is as simple as shown in Figure 1, there are certain requirements for the diagram to be well defined. They are

i. $Z_o + Z_o^*$ is 1-1
ii. $Z_o + Z_L$ is 1-1
iii. $V_g \in R(Z_o + Z_o^*) \cap R(Z_o + Z_L)$

From the diagram, we have

$$V = (Z_0 + Z_0^*)\, I_i = (Z_0 + Z_L)\, I_a$$
$$= V_i + Z_0\, I_i \qquad = V_a + Z_0\, I_a.$$

Define $V_r = V_a - V_i$ and $I_r = -(I_a - I_i)$, then we have $V_r = Z_0\, I_r$.

Define $I_r = S^I\, I_i$, where S^I is called the current-basis scattering operator, then

$$I_r = I_i - I_a = I_i - (Z_0 + Z_L)^{-L}\, (Z_0 + Z_0^*)\, I_i$$

so

$$S^I = I_B - (Z_0 + Z_L)^{-L}\, (Z_0 + Z_0^*), \text{ where } I_B \text{ is the identity mapping on current}$$

space B. Now let K and T be the factorizations of $1/2\cdot(Z_0 + Z_0^*)$ as defined in Thm.IV.2 and Thm.IV.3, i.e.

$$M = 1/2\cdot(Z_0 + Z_0^*) = K\, K^* = T^*\, T.$$

Define $a = K^*\, I_i$, $b = T\, I_r$ and $b = S\, a$, where S is the so-called sacttering operator. Then $I_r = S^I\, I_i$ implies that

$$T^{-R}\, b = S^I\, (K^*)^{-R}\, a, \text{ so } b = T\, S^I\, (K^*)^{-R}\, a. \text{ Hence}$$

$$S = T\, S^I\, (K^*)^{-R} = T\, (K^*)^{-R} - 2\, T\, (Z_0 + Z_L)^{-L}\, K$$
$$= C - 2\, T\, Y_a\, K, \text{ where } C = T\, (K^*)^{-R} \text{ and } Y_a = (Z_0 + Z_L)^{-L}.$$

In order to have a causal scattering operator S, we need a causal C. However, $C = T\, (K^*)^{-R}$ is not causal in general. By Thm.IV.2 and Thm.IV.3,

$C = W\, T_0\, (K_0^*)^{-R}\, U$, where K , T denote the left- and right-factorization repecatively. Therefore, the requirement for the selection of U & W is to make C causal.

Similarly, consider the following network, an n-port network parallel loaded by an n-port network.

a. parallel loaded n-port b. optimal matching situation

Figure 2.

In Figure 2 we have

i. $I_a, I_g, I_i \in B$ = reflexive Banach space.

ii. V_i, V_a B^*

iii. $Y_L, Y_o : (B^*, F^*) \to (B, F)$, linear and bounded.

As before, certain requirements are needed for the circuit diagram to make sence. They are

i. $Y_o + Y_L$ is 1-1

ii. $Y_o + Y_o{}^*$ is 1-1

iii. $I_g \in R(Y_o + Y_o{}^*) \cap R(Y_o + Y_L)$.

With the help of the circuit diagram, the following equations can be easily verified.

$$I_g = (Y_o + Y_o{}^*)\ V_i = (Y_o + Y_L)\ V_o$$

$$= I_i + Y_o\ V_i \quad = I_a + Y_o\ V_a.$$

$$I_r \stackrel{def}{=} -(I_a - I_i),\ V_r \stackrel{def}{=} V_a - V_i.$$

$$I_r = Y_o\ V_r.$$

$V_r \stackrel{def}{=} S^V\ V_i$, where S^V is the so-called voltage-basis

scattering operator.

$S^V = (Y_L + Y_o)^{-L}\ (Y_o{}^* - Y_L) = -I_{B^*} + Z_a\ (Y_o + Y_o{}^*)$, where

$Z_a = (Y_L + Y_o)^{-L}$.

$a \stackrel{def}{=} Q^*\ V_i,\ b \stackrel{def}{=} P\ V_r$, where P, Q are the factorizations of

$1/2\ (Y_o + Y_o{}^*) = P^*\ P = Q\ Q^*$ as mentioned in Thm.IV.2 and Thm.IV.3.

$b \stackrel{def}{=} S\ a$, where S is the scattering operator.

$S = -P\ (Q^*)^{-R} + 2\ P\ Z_a\ Q = -D + 2\ P\ Z_a\ Q$, where $D = P\ (Q^*)^{-R}$.

Similarly, in order to have a causal S, we must have a causal D. However, $D = P\,(Q*)^{-R} = W'\,P_0\,(Q_0*)^{-R}\,U'$, where P_0, Q_0 is the right- and left-factorization of $1/2\cdot(Y_0 + Y_0*)$. Therefore the requirement for the selection of W' & U' is to make D causal.

One of the most useful properties of scattering variables is that they give a measure of the optimal transducer power gain. To see that this property still holds for our generalized scattering operators, let us consider the "power" entering the load network. For the series loaded network,

$$I_a = -I_r + I_i = -\,T^{-R}\,b + (K*)^{-R}\,a$$

$$V_a = V_r + V_i = Z_0\,T^{-R}\,b + Z_0*\,(K*)^{-R}\,a.$$

So the power entering the load is given by:

$$
\begin{aligned}
(I_a, V_a)_B &= ((K*)^{-R}\,a - T^{-R}\,b,\; Z_0\,T^{-R}\,b + Z_0*\,(K*)^{-R}\,a)_B \\
&= ((K*)^{-R}\,a,\; Z_0*\,(K*)^{-R}\,a)_B - (T^{-R}\,b,\; Z_0\,T^{-R}\,b)_B \\
&\quad + ((K*)^{-R}\,a,\; Z_0\,T^{-R}\,b)_B - (T^{-R}\,b,\; Z_0*\,(K*)^{-R}\,a)_B \\
&= (a,\; K^{-L}\,Z_0*\,(K*)^{-R}\,a)_H - (b,\; (T*)^{-L}\,Z_0\,T^{-R}\,b)_H \\
&\quad + (a,\; K^{-L}\,Z_0\,T^{-R}\,b)_H - (K^{-L}\,Z_0\,T^{-R}\,b,\; a)_H \\
&= 1/2(a,\; K^{-L}\,Z_0*\,(K*)^{-R}\,a)_H + 1/2(K^{-L}\,Z_0*\,(K*)^{-R}\,a,\; a)_H \\
&\quad - 1/2(b,\; (T*)^{-L}\,Z_0\,T^{-R}\,b)_H - 1/2((T*)^{-L}\,Z_0\,T^{-R}\,b,\; b)_H \\
&= (a,\; 1/2\cdot K^{-L}\,(Z_0* + Z_0)\,(K*)^{-R}\,a)_H \\
&\qquad - (b,\; 1/2\cdot(T*)^{-L}\,(Z_0* + Z_0)\,T^{-R}\,b)_H \\
&= (a,\; K^{-L}\,K\,K*\,(K*)^{-R}\,a)_H - (b,\; (T*)^{-L}\,T*\,T\,T^{-R}\,b)_H \\
&= (a, a)_H - (b, b)_H \\
&= (a, a)_H - (S\,a,\; S\,a)_H
\end{aligned}
$$

$$= (a, a)_H - (a, S^* S a)_H$$

$$= (a, (I_H - S^* S) a)_H.$$

The above equation indicates that for passive network, i.e.

$$(I_a, V_a)_H \geq 0,$$

$I_H - S^* S$ must be a positive operator. And $S^* S = I_H$ for a lossless network.

The same result can be obtained for the parallel network.

Now let us study an example.

Example 4. i. Let $p, q \in R^+$, $p \geq q$ and $\frac{1}{p} + \frac{1}{q} = 1$.

ii. Let ℓ_p be the current space, ℓ_q be the voltage space. We have $\ell p \subset \ell q$.

iii. Let I_{pq} be the embedding mapping from ℓ_p to ℓ_q.

Then we have:

1. For all $x, y \in \ell p$,

$$(x, I_{pq}\, y) = (x, y) = \sum_{i=1}^{\infty} x_i\, y_i = (y, x)$$

$$\underline{\text{def}}\ (y, I_{pq}{}^* x).$$

Hence, $I_{pq}{}^* = I_{pq}$.

2. I_{pq} is positive since

$$(x, I_{pq}\, x) = (x, x) = \sum_{i=1}^{\infty} x_i^2 \geq 0, \text{ for all } x \in \ell_p.$$

3. Let I_{p2} denote the embedding mapping from ℓ_2 to ℓ_2 and I_{2q} the embedding mapping from ℓ_2 to ℓ_q. Then

$$I_{pq} = I_{2q}\, I_{p2}$$

This relation can be explained by the following diagram.

$$\begin{tikzcd} & \ell_2 \arrow[dr, "I_{2q}"] & \\ \ell_p \arrow[ur, "I_{p2}"] \arrow[rr, "I_{pq}"'] & & \ell_q \end{tikzcd}$$

4. Since we have

$$I_{p2} : \ell_p \to \ell_2 \text{, then}$$

$$I_{p2}^* : \ell_2 \to \ell_p^* = \ell_q.$$

For all $x \in \ell_p$, $y \in \ell_2$, we have

$$(y, I_{p2}\, x)_2 = (y, x)_2$$

$$= \sum_{i=1}^{\infty} y_i\, x_i$$

$$= (x, y)_p$$

Hence,

$$I_{p2}^*\, y = y, \text{ for all } y \in \ell_2, \quad \text{i.e.,}$$

$$I_{p2}^* = I_{2q}.$$

Similarly,

$$I_{2q}^* = I_{p2}.$$

Therefore,

$$I_{pq} = I_{2q}\, I_{p2} = I_{2q}\, I_{2q}^* = I_{p2}^*\, I_{p2}.$$

5. It can easily be proven that I_{p2} and I_{2q} are causal and anti-causal in the usual time structure of the natural numbers. We also have

$$M \;\underline{\text{def}}\; \frac{1}{2}(I_{pq} + I_{pq}) = I_{pq}$$

$$= I_{p2}^*\, I_{p2} = I_{2q}\, I_{2q}^*.$$

6. Using the formula derived for scattering operator with $T = I_{p2}$ and $K = I_{2q}$, we have

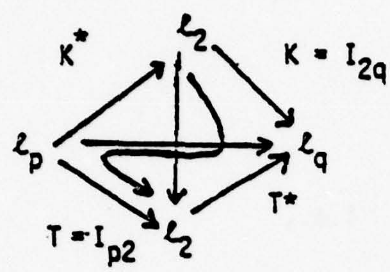
$$S = I_{p2}\,(I_{2q}^*)^{-R} - 2I_{p2}\,(I_{pq} + Z_L)^{-L}\, I_{2q}$$

$$= I_{p2}\, I_{p2}^{-R} - 2\, I_{p2}\, (I_{pq} + Z_L)^{-L}\, I_{2q}$$

$$= \quad I_2 - 2\, I_{p2}\, (I_{pq} + Z_L)^{-L}\, I_{2q}$$

where I_2 is the identity mapping on ℓ_2.

The transformation involved can better be explained by the following diagram:



The scattering operator is meaningful when

$$V_g \in R(I_{pq} + Z_L) \cap R(I_{pq})$$

and

$$I_{pq} + Z_L \text{ is 1-1.}$$

Note in this example operator I_{pq} is not bounded. However, boundedness has nothing to do with the derivation of the scattering operator. Boundedness only guarantees the existance of the factorization. Once the factorization is given, the derivation for the scattering operator follows through accordingly.

VI. Conclusions

The research herein originated with a discussion between one of the authors and Professor Harley Flanders concerning the underlying mathematical nature of electric networks. It was observed that in the scattering formalism the energy disapation of a network was given in terms of the norm of the network variables by

$$E = ||a||^2 - ||b||^2$$

whereas, in the immittance formalisms the energy was given by the inner product (or functional) equality

$$E = (v,i)$$

Since a network is fundamentally an energy processing system, one might initially interprete these equalities as implying that the scattering variables are naturally defined in a Banach space since only the norm is required to define their energy, whereas, the immittance variables for which an inner product is required to define energy, are naturally defined in Hilbert space. In fact, the situation is just the contrary. The immittance variables may naturally be extended to Banach space by working simultaneously with B and B* whereas, the development of the present paper indicates that the scattering variable "live" in a Hilbert space even when their corresponding immittance variables are defined in Banach space. Indeed, we believe that the primary contribution of the present work is the observation that certain problems naturally "live" in Hilbert space. Moreover, they may be transformed into a Hilbert space even when initially defined in Banach space.

APPENDIX

1. Proof of Thm.III.2:

A. i. By Thm.III.1, there exists a Hilbert space $\hat{H}$ and a linear bounded operator $\hat{K}$ from $\hat{H}$ to B^* such that $Q = \hat{K}\hat{K}^*$.

ii. Define $H = \overline{R(\hat{K}^*)}$ and $K = \hat{K}|_H : H \to B^*$, then $K^*: B^{**}=B \to H$.

iii. $(K^* b, x)_H = (b, K x)_B = (b, \hat{K} x)_B = (\hat{K}^* b, x)_H$, for all $b \in B$, $x \in H$.

Since $K^* b$, $\hat{K}^* b \in H$, so $K^* b = \hat{K}^* b$, for all $b \in B$.

iv. $K K^* b = K \hat{K}^* b = \hat{K}\hat{K}^* b = Q b$, for all $b \in B$. So $K K^* = \hat{K}\hat{K}^* = Q$.

v. Since $H = \overline{\hat{K}^*[B]} = \overline{K^*[B]}$, so K^* has dense range. So K is 1-1.

B. i. Define $H^t = \overline{R(K^* F^t)}$ and let E^t be the orthogonal projection on H. Then $(E^t)^2 = (E^t)^* = E^t$.

ii. Since H^t becomes $\overline{R(K^*)}$ which is H, as $t \to \infty$, so $\lim E^t = I_H$ ($E^t \to I_H$ weakly).

iii. When $s \leq t$,

$H^s = \overline{R(K^* F^s)}$, $H^t = \overline{R(K^* F^t)}$ and $R(F^w) \subseteq R(f^t)$.

Since $F^t b = F^s b + F[s,t) b$, for all $b \in B$, and for $b' \in R(F^s)$, $b' = F^s b'$, so

$F^t b' = F^t (F^s b') + F[s,t) (F^s b') = F^s b' = b'$,

i.e. $b' \in R(F^t)$. So $\overline{R(K^* F^s)} \subseteq \overline{R(K^* F^t)}$ and

$H^s = \overline{R(K^* F^s)} \subseteq \overline{R(K^* F^t)} = H^t$. So $E^t E^s = E^s E^t = E^s$.

iv. With E defined for all $(-\infty, t)$, $t \in R$, E can be extended to $\mathcal{B}(R)$ uniquely to be a spectral measure, i.e. a resolution of identity.

v. Since $E^t[H] = H^t = \overline{R(K^* F^t)}$, by definition, so K^* is a right-maxiphase (Def.II.5). So K is a left-miniphase (Thm.II.8).

C. i. Let $\tilde{K}: (\tilde{H}, \tilde{E}) \to (B^*, F^*)$ be another left-miniphase factorization of Q.

ii. Define W on $R(\tilde{K}^*)$ by $W(\tilde{K}^* b) = K^* b$. W is well-defined, for if $\tilde{K}^* b = \tilde{K}^* a$, then

$K K^* a = Q a = \tilde{K} \tilde{K}^* a = \tilde{K} \tilde{K}^* b = Q b = K K^* b$. But K is 1-1, so $K^* a = K^* b$.

iii. For any $b \in B$,

$$\|W \tilde{K}^* b\|_H^2 = \|K^* b\|_H^2 = (K^* b, K^* b)_H$$
$$= (b, K K^* b)_B = (b, \tilde{K} \tilde{K}^* b)_B$$
$$= (\tilde{K}^* b, \tilde{K}^* b)_{\tilde{H}} = \|\tilde{K}^* b\|_{\tilde{H}}^2,$$

so W is isometric on $R(\tilde{K}^*)$. Since $R(\tilde{K}^*)$ is dense in $\tilde{H}$, W can be ismetrically extended to $\tilde{H}$.

iv. For all $z \in R(\tilde{K}^*)$, $z = \tilde{K}^* x$, for some $x \in B$, hence

$K W z = K W \tilde{K}^* x = K K^* x = \tilde{K} \tilde{K}^* x = \tilde{K} z$. So

$K W = \tilde{K}$ over $R(\tilde{K}^*)$. So $K W = \tilde{K}$ over H via continuity as W extended to H.

v. By the same argument, there exists $V: H \to \tilde{H}$, V isometric, such that $\tilde{K} V = K$. So $K W V = K$, and hence

$W V = I_H$, since K is 1-1. Similarly, $V W = I_{\tilde{H}}$.

vi. With $K W = \tilde{K}$, we have

$$(F^t)^* K E^t W = (F^t)^* K W = (F^t)^* \tilde{K} = (F^t)^* \tilde{K} \tilde{E}^t$$
$$= (F^t)^* K W \tilde{E}^t = (F^t)^* K E^t W \tilde{E}^t,$$

(K, $\tilde{K}$ are causal). So

$(F^t)^* K (E^t W - E^t W \tilde{E}^t) z = 0$ for all $z \in H$, hence

$E^t (E^t W - E^t W \tilde{E}^t) z = 0$, ($K$ is left-miniphase),

i.e. $E^t W = E^t W \tilde{E}^t$, so W is causal.

vii. Similarly, V is causal. But $W = V^*$ which is anti-causal (Thm.II.6), so W is memoryless.

2. Proof of Thm.III.3.

Define $T = K^*$, then $T^* = K$, and define $\underset{\sim}{H}_t = \overline{R(T F_t)}$. Then the rest of the proof follows as in Thm.III.2.

References:

1. Saeks, R., Resolution Space, Operators and Systems, New York, Springer-Verlag, 1973.

2. Tung, L.J., "Random Variables, Wiener-Hopf Filtering and Control Formulated in Abstract Spaces," Ph.D. Thesis, Texas Tech University, 1977.

3. Parthasarathy, K.R., Probability Measures on Metric Spaces, Academic Press, 1967.

4. Balakrishnan, A.V., Introduction to Optimization Theory in a Hilbert Space, New York, Springer-Verlag, 1971.

5. Bachman, G. and Narici, L., Functional Analysis, New York, Academic Press, 1966.

6. Chobanian, S.A., "On a Class of Functions of a Banach Space Valued Stationary Stochastic Process", Sakharth SSR Mecn. Acad. Moambe 55, pp. 21-24, 1969.

7. Chobanian, S.A., "On Some Properties of Positive Operator-Valued Measures in Banach Spaces", Sakharth SSR Mecn. Acad. Moambe 57, pp. 273-276, 1970.

8. Vakhania, N.N., "The Covariance of Random Elements in a Banach Space", Thbilis Sanelmc. Univ. Gamoqeneb. Math. Inst. 2 (1969), pp. 179-184.

9. Masani, P., "An Explicit Treatment of Dilation Theory", unpublished notes.

10. Rohrer, R.A., "The Scattering Matrix: Normalized to Complex n-port Load Networks", IEEE Trans. on Circuit Theory, Vol. CT-12, pp. 223-230, 1965.

11. Luenberger, D.G., Optimization by Vector Space Methods, New York, J. Wiley and Sons, 1969.

12. Saeks, R., "Reproducing Kernel Resolution Space and its Applications", in Journal of the Franklin Institute, Vol. 302, No. 4, pp. 331-355, Oct. 1976.

13. Kailath, T. and Duttweiler, D., "An RKHS Approach to Detection and Estimation Problems-Part III: Generalized Innovations and a Likelihood-ratio Formula", IEEE Trans. Info. Thy., Vol. IT-18, pp. 730-745, 1972.

14. Duttweitler, D., "Reproducing Kernel Hilbert Space Techniques for Detection and Estimation Problems", Ph.D. Thesis, Stanford Univ., 1970.

15. Grenander, U., Probabilities on Algebraic Structures, Wiley, New York, 1963.

16. Aronazajn, N., "Theory of Reproducing Kernels", Trans. on the AMS, Vol. 63, pp. 337-404, 1950.

17. Porter, W.A., "Some Circuit Theory Concepts Revisited", Int. J. on Control, Vol. 12, pp. 443-448, 1970.

18. Schnure, W.K., "Controllability and Observability Criteria for Causal Operators on a Hilbert Resolution Space", Proc. of the 14th MSCT, Univ. of Denver, 1971.

19. Saeks, R., "The Factorization Problem - A Survey", Proc. of IEEE, Vol. 64, No. 1, Jan. 1976.

STABILITY AND HOMOTOPY

R. Saeks and R.A. DeCarlo

*This research supported in part by AFOSR grant 74-2631.

STABILITY AND HOMOTOPY*

R. Saeks
Texas Tech Univ.
Lubbock, Tx. 79409

R. A. DeCarlo
Purdue Univ.
W. Lafayette, In. 47907

ABSTRACT

A generalization of the classical Nyquist stability criterion to nonlinear and time-varying systems is obtained via an appropriate homotopy argument in the space of causal invertible (possibly nonlinear) operators. Although the resulting stability test is only a sufficient condition in its most general form it reduces to the classical necessary and sufficient Nyquist criterion for linear time-variant systems characterized by a transfer function or transfer function matrix.

Although appearently abstract the homotopic nature of the proof proves to be quite transpearent and, as such, many of the classical sufficient conditions for non-linear or time-varying systems can be derived from the generalized Nyquist criterion by simply constructing a homotopy (continous deformation) from the given system to a ststem which is known to satisfy the generalized Nyquist criterion. This is illustrated via a simple derivation of the Circle criterion as a corollary to the generalized Nyquist criterion.

INTRODUCTION

When one discusses alternatives in multivariable control the classical debate between the advocates of frequency and time domain techniques usually comes to the fore. The former is highly intuitive but restricted to linear time-invariant systems whereas the latter is amenable to efficient computational procedures and is readily extendable to nonlinear and time-varying systems. A third alternative is the operator theoretic approach wherein the system is modeled by an operator on Hilbert space. In the view of the author such an approach to the control problem achieves the best of both the time and frequency domain techniques. Since the operator theoretic model is defined in the time domain the resultant control techniques often hold for nonlinear and time-varying systems. On the other hand, operator theoretic techniques are formally quite similar to the operational calculus associated with the frequency domain. As such, the intuitive character of frequency domain control theory often carries over to the operator theoretic approach.

*This research supported in part by AFOSR Grant 74-2631.

The purpose of the present paper is to illustrate the potential of the operator theoretic approach to multivariable control via the derivation of a generalized Nyquist criterion which is applicable to nonlinear and time-varying systems modeled by finite gain operators on Hilbert space. Although only a sufficient condition, in general, the technique reduces to the classical necessary and sifficient Nyquist criterion for linear time-invariant multivariable systems[1,2] and it appears to be "tight" in the general case.

Although the derivation holds in an abstract Hilbert Resolution Space[3] for the sake of brevity the present discussion will be restricted to the case of systems defined on the space L_2^n composed of n-vectors of square integrable functions. For this space we define the norm

1. $$||f||^2 = \int_R f(q)^t f(q)\, dq$$

and a family of <u>truncation operators</u> $P^t:L_2^n \longrightarrow L_2^n$ by

2. $$(P^t f)(q) = \begin{cases} f(q) & q \leq t \\ 0 & q > t \end{cases}$$

An operator $T:L_2^n \longrightarrow L_2^n$ is said to have <u>finite gain</u>[1] if there exist constants M and N such that

3. $$||Tf|| \leq M||f|| + N$$

for all f in L_2^n. In some sence the constant M plays the role of a norm for the nonlinear operator T. Of course, for linear operators M may be taken to be the norm of T with N = 0 in which case T has finite gain if and only if it is bounded. In the nonlinear case if an operator has a finite Lischitz constant then it is also a finite gain operator though many finite gain operators do not admit Lipschitz constants.[3]

We say that an operator $T:L_2^n \longrightarrow L_2^n$ is causal[3] if

4. $$P^t T = P^t T P^t$$

for all t. It is easily shown[3] that the causal operators are closed under operator addition and multiplication and limits taken in the topology defined by the gain constants M and N. Unfortunately, they are not closed under the operation of operator inversion. A classical example of this is the unit delay whose inverse is a predictor. Indeed, the question of determining whether or not T^{-1} is causal from the properties of T is completely equivalent to the question of determining whether or not a feedback system is stable from the properties of its open loop gain[1,3,4,5,6]. As such, the following discussion of the generalized Nyquist criterion will be formulated in terms of the problem of determining whether or not the inverse of a causal operator is causal, the solution to the feedback system stability problem being obtained by applying these results to the return difference operator.[3]

THE NYQUIST CRITERION

The classical Nyquist criterion is usually formulated in terms of the degree of the system frequency response. For such frequency responses, however, their degree is simply a representation of their homotopy class[7,8] and hence we formulate the present discussion in terms of homotopic operators. In particular, we say that operators T_o and T_1 are *homotopic* in the space of causal invertible operators*, C(0), if there exists a continous operator valued function $T:I\ C(0)$ mapping the interval [0,1] to the group of causal invertible operators such that $T(0) = T_o$ and $T(1) = T_1$. Our main theorem now may be stated as:

> *THEOREM*: Let T_o and T_1 be fininte gain operators which are homotopic in C(0). Then if T_o has a causal inverse so does T_1.

The proof is based on the following lemma usually known as the *small gain theorem*[1,3].

*Both the operators and their inverses are assumed to be finite gain but the inverses need not be causal.

Lemma: Let T be finite gain causal operator for which $M < 1$. Then, if the operator $(1 + T)$ has a finite gain inverse, the inverse is causal.

A proof of the lemma appears in reference 1 and will not be repeated here.

Proof of the Theorem: Let T be a homotopy from T_0 to T_1 and assume that $T(t_0)$ has a causal inverse. Now let $|t - t_0| < \varepsilon$ and write

5. $$T(t) = T(t_0) + (T(t)-T(t_0)) = [1 + (T(t)-T(t_0))T(t_0)^{-1}]\, T(t_0)$$

By hypothesis $(T(t)-T(t_0))T(t_0)^{-1}$ is causal and has a gain constant $M < 1$ if ε is chosen sufficiently small (by continuity and the fact that $T(t_0)^{-1}$ if finite gain). Moreover,

6. $$[1 + (T(t)-T(t_0))T(t_0)^{-1}]^{-1} = T(t_0)T(t)^{-1}$$

exists and is finite gain since $T(t_0)$ is finite gain and $T(t)$ has a finite gain inverse by hypothesis. As such, the small gain theorem implies that

7. $$T(t)^{-1} = T(t_0)^{-1}\, [1 + (T(t)-T(t_0))T(t_0)^{-1}]^{-1}$$

is causal since it is the product of two causal operators. Finally, since the [0,1] interval is a compact set one can piece together finitely many ε -intervals to show that $T(1)^{-1} = T_1^{-1}$ is causal if $T(0)^{-1} = T_0^{-1}$ is causal, thereby completeing the proof.

APPLICATIONS

Intuitively the theorem states that the property of a finite gain causal operator having a finite gain causal inverse is an invariant of the arcwise connected[8] component of C(0) in which the operator lies. To obtain a test for causal invertibility it therefore suffices to show that a given operator lies in the same arcwise connected component of C(0) as an operator which is known to admit a causal inverse. In particular, the following corollary reduces to the classical Nyquist condition for linear time inveriant multivariable systems.

Corrollary: Let T be finite gain operator which is homotopic to the identity operator in C(0). Then T has a causal inverse.

Of course, the identity in the above corollary could equally well be replaced by any operator satisfying one of the classical sufficient conditions for causal invertibility[1,3]; operators satisfying the conditions of the small gain theorem[1,3], monotonic operators[3], operators of the form 1 + S where S is strictly causal[3,7], etc. Interestingly, however, each of these classes of operators lie in the same arcwise connected component as the identity and the fact that they admit causal inverses is most easily derived from the above corollary rather than conversely. Another such class of operators which fall in the same arcwise connected component as the identity are the causal operators for which zero lies in the unbounded component of their resolvant set. Indeed, the proof that such operators admit causal inverses appearing in reference 9 is almost identical to the proof of the present theorem but with a restricted class of homotopy. In fact, the present theorem is a simple extension of the earlier result though considerably tighter. In particular, the result of reference 9 is not necessary and sufficient in the linear time-invariant case and assumes that the operators involved admit finite Lipschitz constants.

An alternative way of looking at the above theorem is as a perturbation theorem wherein large purturbations are allowed so long as they are continous relative to the operator topology defined by the operator gain. As such, it should not be surprising that many of the small perturbation results of classical stability theory follow from the generalized Nyquist criterion. For instance, one may derive the circle criterion[1] via a two step homotopy. First, one deforms the nonlinear term to a linear (lying in the middle of the sector associated with the nonlinearity) and then one deforms the resultant linear system into the identity operator via the classical Nyquist criterion. Here, a combination of the sectoral bound[1] and the requirement that the spectrum of the linear part of the system lie outside of an appropriate disk suffices to assure that the operator lies in C(0) at every point in the homotopy and hence justifies the application of the Theorem.

Ofcourse, once this homoptopic point of view is adopted, numerous generalizations become appearent.

CONCLUSIONS

The purpose of this short paper has been two-fold. First, we believe that the generalization of the Nyquist criterion presented may prove to be an extremely powerful tool of stability theory. Indeed, we conjecture that this single elementary results subsumes most, if not all, of classical stability theory. Secondly, however, we believe that it illustrates the power of operator theoretic techniques in control which have the potential of achieving the best of both the time and frequency domain worlds. Indeed, such techniques yield natural and intuitive generalizations of the classical frequency domain concepts without the linearity and time-invariance restrictions usually associated therewith.

REFERENCES

1. Desoer, C.A., and M. Vidyasager, Feedback Systems: Input-Output Properties, New York, Academic Press, 1975.

2. Saeks, R., and R.A. DeCarlo, Interconnected Dynamical Systems, New York, Marcel-Dekker, (in press).

3. Saeks, R., Resolution Space, Operators, and Systems, Heidelberg, Springer-Verlag, 1973.

4. Willems, J.C., "Stability, Instability, Invertibility, and Causality", SIAM Jour. on Cont., Vol. 7, pp. 645-671, (1969).

5. DeSantis, R.M., Ph.D. Thesis, Univ. of Michigan, 1971.

7. DeCarlo, R.A., and R. Saeks, "The Encirclement Condition: An Approach Using Algebraic Topology", Int. Jour. on Cont., (to appear).

8. Massey, W.S., Algebraic Topology: An Introduction, New York, Harcourt, Brace and World, 1967.

9. Saeks, R., "On the Encirclement Condition and its Generalization", IEEE Trans. on Circuits and Systems, Vol. CAS-22, pp. 780-785, (Oct. 1975).

VARIATIONS ON THE NONLINEAR NYQUIST CRITERIA*

R.A. DeCarlo and R. Saeks

*This research supported in part by AFOSR grant 74-2631.

VARIATIONS ON THE NONLINEAR NYQUIST CRITERIA*

R. A. DeCarlo and R. Saeks
Texas Tech University
Lubbock, Texas

ABSTRACT

Classically the study of closed loop system stability is approached through frequency domain techniques, e.g. the Nyquist and Hurwitz criteria. In the nonlinear case frequency response is not well defined; however, one of the authors has recently shown that the spectrum of a nonlinear operator can be used in lieu of the usual Nyquist plot as a means of generalizing the Nyquist criteria to the nonlinear case.

Through some perturbation techniques we characterize in this paper the stability of nonlinear operators by the more accessible "approximate point spectrum" as opposed to the entire spectrum.

I. *INTRODUCTION*

Recently one of the authors demonstrated that the stability of a closed loop system rests squarely on knowledge of the spectrum of the operator which represents the open loop gain. (1) The system may be nonlinear, multivariable, and/or time-varying. For a linear operator representing the open loop gain the spectrum and frequency response coincide, however, computation of the spectrum of a nonlinear operator is not, in general, a trivial exercise. (1) (2)

This paper shows that knowledge of the more easily computed approximate point spectrum is adequate to answer the stability question. (3) In reference 1, it is shown that if the spectrum of the operator (representing the open loop gain in a unity feedback system) does not encircle "-1" then the system is stable. Essentially, this is equivalent to the requirement that the component of the resolvant (the complement of the spectrum) contain the point "-1"--i.e., "-1" is not disconnected from infinity by the spectrum. It is shown here that the infinite component of the resolvant and the infinite component of the complement of the approximate point spectrum are identical. Thus if the approximate point spectrum of the aforementioned operator does not encircle "-1", then the system is stable.

Finally the set of complex numbers $\{\lambda = \hat{y}(\omega)/\hat{x}(\omega)\}$, where $\hat{x}(\omega)$ and $\hat{y}(\omega)$ are the Fourier transforms of the input and output respectively, is shown to contain the approximate point spectrum.

*This research supported in part by Air Force Office of Scientific Research Grant AFOSR 74-2631.

In some cases this set is very large and can be the whole complex plane. This objection is offset somewhat by the fact that the spectrum of trivial nonlinear operators, such as a squarer, may also be very large.

II. THE APPROXIMATE POINT SPECTRUM AND STABILITY

All operators map L_2^n to itself unless otherwise specified. An operator W is causal if whenever

$$x(t) = y(t) \quad t < T; \; f,g \text{ in } L_2^n \tag{1}$$

for some T, then

$$(Wx)(t) = (Wg)(t) \qquad t < T. \tag{2}$$

The norm of W, $||W||$, is the usual Lipschitz norm. W is stable if it is both causal and bounded. (4) (5) The spectrum of an operator, W (possibly nonlinear), is the set of complex numbers λ, such that the operator $(\lambda - W)$ does not have a bounded inverse.* $\sigma(W)$ denotes the spectrum of the operator W. $\sigma(W)$ is a compact set. The resolvant set of W is $\rho(W)$ which is the complement of the spectrum of W in the complex plane. Clearly $\rho(W)$ is open.

The following is a recent theorem by one of the authors. (1)

<u>THEOREM 1</u>: Let the open loop gain of a (possibly) nonlinear feedback system be represented by a stable *unbiased transformation* K, *mapping* L_2^n to itself. Then if the spectrum of K in the algebra of Lipschitz continuous unbiased operators does not encircle the point "-1", the feedback system is stable.

The theorem says that the closed loop system is stable if "-1" is in the infinite component of the resolvant. For the case of a single input single output, linear, time invariant system whose open loop gain is characterized by the frequency response $\hat{H}(\omega)$ the Nyquist plot for $\hat{H}(\omega)$ is precisely the spectrum of the open loop gain. Hence the above theorem coincides with the classical Nyquist test.

Typically the spectrum of a nonlinear operator W is difficult to compute. A characterization of stability using the approximate point spectrum offers a more accessible route, at least theoretically. To this end we denote the approximate point spectrum of the (possibly) nonlinear operator W as $\pi(W)$. $\pi(W)$ is the set of all complex numbers, λ, such that for all $\varepsilon > 0$, there exists an $x \neq 0$, such that $||(\lambda - W)x|| \leq \varepsilon \, ||x||$. $\pi(W)$ is a closed set and contains the point spectrum (the set of complex λ, such that there exists x satisfying $Wx = \lambda x$). We denote the complement of $\pi(W)$ by $\gamma(W)$. A complex number, γ, is in $\gamma(W)$ if

*The symbol $(\lambda - W)$ where λ is a scalar is used to denote the operator $(\lambda I - W)$ where I is the identity operator.

there exists an $\varepsilon > 0$, such that $||(\lambda - W)x|| > \varepsilon||x||$, for all $x \neq 0$. In proving theorems this definition seems to have more utility as in the following proposition.

Proposition: $\pi(W) \subset \sigma(W)$.

Proof: Suppose $\lambda \notin \sigma(W)$. We show $\lambda \notin \pi(W)$ implying $\pi(W) \subset \sigma(W)$. Since $\lambda \notin \sigma(W)$ we have $(\lambda - W)^{-1}$ exists and is bounded. With this fact consider the norm of x.

$$||x|| = ||(\lambda - W)^{-1}(\lambda - W)x|| \leq ||(\lambda - W)^{-1}|| \; ||(\lambda - W)x|| \tag{3}$$

Setting $\varepsilon = 1/||(\lambda - W)^{-1}||$, we conclude that

$$||(\lambda - W)x|| \geq \varepsilon||x||. \tag{4}$$

Thus $\lambda \notin \pi(W)$ as was to be shown.

Since $\pi(W) \subset \sigma(W)$ we have $\rho(W) \subset \gamma(W)$. $\gamma(W)$ and $\rho(W)$ are open sets since they are the complements of closed sets. As with any set, both are the union of their connected components. Both contain a unique infinite component. Necessarily, the infinite component of $\gamma(W)$ contains the infinite component of $\rho(W)$. A corollary to the following lemma shows that these infinite components are, in fact, identical.

Lemma 1: Let W be a (possibly) nonlinear operator. Let γ_a be a connected component of $\gamma(W)$. If γ_a contains a point in $\rho(W)$, then $\gamma_a \subset \rho(W)$.

Proof: Essentially we show that each connected component of $\rho(W)$ coincides with a connected component of $\gamma(W)$. Let γ_a be a connected component of $\gamma(W)$. Suppose a point p is an element of both γ_a and $\rho(W)$. Suppose further that q is any other point in γ_a and that ℓ is a path connecting the points p and q.

Since p is in the resolvant, $(p - W)^{-1}$ exists and is bounded. The task is to show that $(q - W)^{-1}$ exists implying $\gamma_a \subset \rho(W)$. Combining this fact with the above proposition, we will have every connected component of $\rho(W)$ coinciding with some connected component of $\gamma(W)$.

The idea of the proof is to use the definition of $\gamma(W)$ and the compactness of the path, ℓ, to find an ε-ball about the point p, such that for any λ in the ε-ball, $(\lambda - W)^{-1}$ exists and is bounded. It turns out that the ε-ball depends only on a single constant. Thus a finite number of ε-balls can be pieced together along the path, ℓ, so that the arbitrary point q is in the resolvant. The details now follow.

By definition, for any λ in γ_a, there exists an $m_\lambda > 0$, such that

$$||(\lambda - W)\, x|| > m_\lambda ||x|| \tag{5}$$

for all x in L_2^n. We can choose m_λ continuously by taking

$$\overline{m}_\lambda = \sup\{m_\lambda\} \tag{6}$$

where the sup is taken over all m_λ satisfying the above inequality. This modifies the inequality to

$$||(\lambda - W)x|| \geq \overline{m}_\lambda ||x||. \tag{7}$$

Since the path ℓ is compact, $\overline{m}_\lambda$ achieves its minimum for some λ_o in ℓ. Define m = $\overline{m}_{\lambda_o} > 0$. Thus for all λ in ℓ we conclude

$$||(\lambda - W)x|| \geq m||x||. \tag{8}$$

Consequently if λ is in ℓ and $(\lambda - W)^{-1}$ exists, then

$$||(\lambda - W)^{-1}|| \leq 1/m. \tag{9}$$

It remains to show that $(\lambda - W)^{-1}$ exists for all λ in ℓ. Define $S_m(p)$ = $\{\lambda/|\lambda - p| < m\}$ to be an m-ball about the point p. Let λ be in $S_m(p)$, then

$$(\lambda - W) = ((\lambda - p) + (p - W)) = ((\lambda - p)(p - W)^{-1} + 1)(p - W). \tag{10}$$

This factorization is valid since $(p - W)^{-1}$ exists and is bounded. The norm

$$||(\lambda - p)(p - W)^{-1}|| \leq |\lambda - p|\ ||(p - W)^{-1}|| < m(1/m) = 1. \tag{11}$$

By the contraction mapping theorem

$$[(\lambda - p)(p - W)^{-1} + 1]^{-1}$$

exists and is bounded. (6) Since $(p - W)^{-1}$ exists and is bounded, the same is true of $(\lambda - W)^{-1}$ for all λ in $S_m(p)$.

All p_o in ℓ, such that $(p_o - W)^{-1}$ exists and is bounded can be enclosed by an ε-ball, $S_\varepsilon(p_o)$, where ε depends only on m, which depends on the path ℓ. Piecing a finite number of ε-balls together we conclude $(q - W)^{-1}$ exists and is bounded. Since $\gamma(W)$ is open, each component of $\gamma(W)$ is open. Thus for any point q in γ_a, there exists a path and a number $m = m(\ell)$, such that every point in γ_a is in $\rho(W)$. Thus the theorem is proved.

<u>Corollary 1</u>: The infinite components of $\rho(W)$ and $\gamma(W)$ are identical.

Corollary 2: $\rho(W)$ is the union of some subset of the set of connected components of $\gamma(W)$.

Proof: Pick all components of $\gamma(W)$ which contain a point in $\rho(W)$. By the above lemma, their components are contained in $\rho(W)$. By the previous proposition the result follows.

Corollary 3: Let W be a bounded operator. If $\gamma(W)$ has only one component, then $\sigma(W) = \pi(W)$.

Proof: Since W is bounded $\gamma(W)$ and $\rho(W)$ contain identical infinite components. Therefore since there is only one component of $\gamma(W)$, $\rho(W) = \gamma(W)$ implying that $\sigma(W) = \pi(W)$.

With these statements as a base, a modified general Nyquist criteria follows.

THEOREM 2: Let the open loop gain of a (possibly) nonlinear feedback system be represented by a stable unbiased transformation, W, mapping L_2^n to itself. Then if the approximate point spectrum of K in the algebra of Lipschitz continuous unbiased operators does not encircle the point "-1", the feedback system is stable.

Proof: We remark that "not encircle -1" is equivalent to "-1" in the infinite component of the resolvant. Thus "-1" is in the infinite component of $\gamma(W)$. Therefore if the approximate point spectrum does not encircle "-1" neither does $\sigma(W)$ and conversely.

III. A COVERING OF THE APPROXIMATE POINT SPECTRUM

For a linear single input single output operator, H, the frequency response $\hat{H}(\omega) = \hat{y}(\omega)/\hat{x}(\omega)$, where $\hat{y}(\omega)$ and $\hat{x}(\omega)$ are the Fourier transforms of the output and input functions respectively, is the spectrum. In the nonlinear case, it appears that the relevent object to study is in fact $\{\hat{y}(\omega)/\hat{x}(\omega)\}$ since it offers a covering of the approximate point spectrum even though $\hat{H}(\omega)$ is undefined.

Let W be a (possibly) nonlinear operator. Suppose $Wx = y$. Define $S(W) = \text{Closure}\,\{\lambda \mid \lambda = \hat{y}(\omega)/\hat{x}(\omega)$ where $\hat{y}(\omega)$ and $\hat{x}(\omega)$ are Fourier transforms$\}$. With these assumptions we have the following lemma.

Lemma 2: $\pi(W) \subset S(W)$.

Proof: Suppose $\lambda \notin S(W)$, then since $S(W)$ is a closed set

$$\left|\lambda - \frac{\hat{y}(\omega)}{\hat{x}(\omega)}\right| > \varepsilon \tag{12}$$

for all $x \neq 0$, for all ω, and some $\varepsilon > 0$. Now by Parsevals equality

$$||(\lambda - W)x||^2 = ||\lambda\hat{x}(\omega) - \hat{y}(\omega)||^2 \tag{13}$$

$$= \int_{-\infty}^{+\infty} |\lambda\hat{x}(\omega) - \hat{y}(\omega)|^2 \, d\omega \tag{14}$$

$$= \int_{-\infty}^{+\infty} |\lambda - \frac{\hat{y}(\omega)}{\hat{x}(\omega)}|^2 |\hat{x}(\omega)|^2 d\omega \tag{15}$$

$$= |\lambda - \frac{\hat{y}(\omega_0)}{\hat{x}(\omega_0)}|^2 \int_{-\infty}^{+\infty} |\hat{x}(\omega)|^2 d\omega \tag{16}$$

by the mean value theorem. It now follows that (16) is greater than $\varepsilon^2||x||$. The lemma is true.

Suppose $Wx = x^2$. Utilizing the global inverse function theorem, one can show the spectrum of W to be the whole complex plane. (7) (8) Thus one could expect the set S(W) to be large. For appropriately restricted weakly additive operators $\hat{y}(\omega)$ is well defined in terms of $\hat{x}(\omega)$ and knowledge of the operator W.

IV. CONCLUSIONS

Apparently the approximate point spectrum is the interesting object of study in the stability question. Moreover knowledge of $\pi(W)$ offers a sufficient condition for when $\sigma(W) = \pi(W)$. Perhaps a variation will offer a necessary and sufficient condition. Lastly the set S(W) covers $\pi(W)$. The set S(W) is intuitively satisfying since it can be interpreted as the frequency gain or frequency response of the system.

V. REFERENCES

1. Saeks, R., "On the Encirclement Criterion and its Generalization", IEEE Transactions on Circuit Theory, October, 1975.

2. Barman, J. F., and J. Katzenelson, "A Generalized Nyquist-type Stability Criterion for Multivariable Feedback Systems", Memo No. ERL-383, Electronics Research Laboratory, University of California at Berkeley, 1973. (Also to appear in the International Journal on Control.)

3. Halmos, P. K., Introduction to Hilbert Space and the Theory of Spectral Multiplicity, New York, Chelsea, 1951.

4. Willems, J. C., "Stability, Instability, Invertibility and Causality", SIAM Journal on Control, Vol. 7, pp. 645-671, (1969).

5. Saeks, R., Resolution Space, Operators and Systems, Heidelberg, Springer-Verlag, 1973.

6. Luenberger, D. G., Optimization by Vector Space Methods, John Wiley and Sons, Inc., New York, 1969.

7. Palais, R. S., "Natural Operations on Differential Forms", Trans. Amer. Math. Soc., Vol. 92, pp. 125-141, 1959.

8. Wu, F. F. and Charles A. Desoer, "Global Inverse Function Theorem", IEEE Transactions on Circuit Theory, March, 1972.

A NEW CHARACTERIZATION OF THE NYQUIST STABILITY CRITERION*

R.A. DeCarlo and R. Saeks

*This research supported in part by AFOSR grant 74-2631.

A NEW CHARACTERIZATION OF THE NYQUIST STABILITY CRITERION*

R. DECARLO & R. SAEKS

ABSTRACT

The usual proof of the Nyquist Theorem depends heavily on the agrument principle. The argument supplies unneeded information in that it counts the number of encirclements of "-1". Stability of a system requires an encirclement or a no-encirclement test. Using homotopy theory, this paper offers a more intuitive approach. We believe this approach will lead to practical generalizations. For example, systems characterized by several complex variables such as multi deminsional digital filters.

I. INTRODUCTION

This paper introduces a characterization of the Nyquist criterion using homotopy theory, a branch of algebraic topology. The authors emphasize the intuition and motivation for this approach. The hope is to aid interested readers to further extend and apply these ideas. In this vein, proofs are omitted so as to simplify the presentation. Details can be found in the references. With this philosophy in mind, let us define the type of system we will be discussing.

As illustrated in Figure 1, let $\hat{g}(s)$ be a rational function in the complex variable s (bounded at $s = \infty$**) representing the open loop gain of scalar single loop feedback system. The closed loop system has transfer function $\hat{h}(s) = \hat{g}(s)/[1+\hat{g}(s)]$. The closed loop system is stable if and only if all poles of $\hat{h}(s)$ are in the open left half plane denoted by $\mathbb{C}_-$ (where $\mathbb{C}$ will denote the entire complex plane).

The Nyquist Criterion states that the closed loop system is stable if and only if the Nyquist plot of $\hat{g}(s)$ (i.e. the image of the Nyquist contour under the map $\hat{g}(\cdot)$) does not encircle nor pass through "-1". If the Nyquist plot passes through "-1" there is a pole on the imaginary axis; if the Nyquist plot encircles "-1", there is a pole in the open right half plane, which we will denote by $\mathbb{C}_+$ ($\mathbb{C}_+$ will denote the closed right

**This boundedness condition can be dispensed with & is added only to ease the exposition.

*Supported in part by AFOSR Grant 74-2631d

half plane). The following section constructs the required machinery of homotopy theory

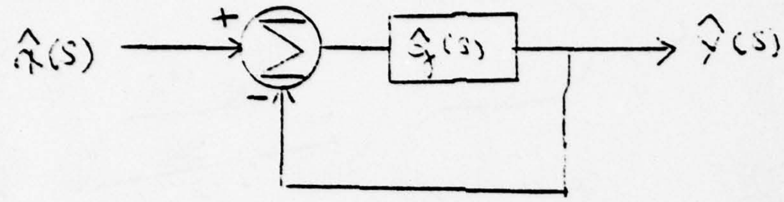


Figure 1

II. MATHEMATICAL PRELIMINARIES & BACKGROUND

Basic to homotopy theory is the concept of a path. A path or a curve in the complex plane is a continuous function of bounded variation (2) $\gamma : [0,1] \to \mathbb{C}$. γ is a closed path if $\gamma(0) = \gamma(1)$. γ is a simple closed path if γ is a closed path and has no self intersections. The image of $I = [0,1]$ under γ is called the trace of γ and is denoted by $\{\gamma\}$.

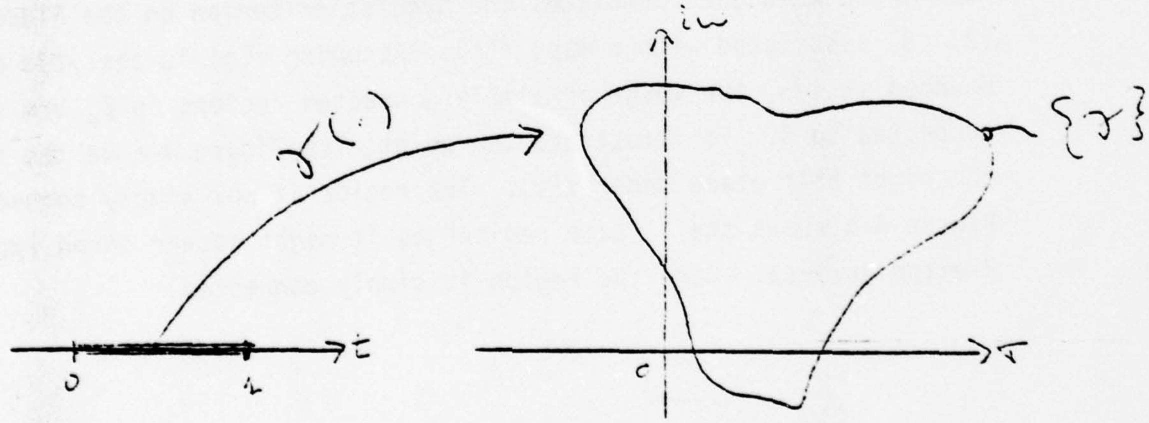


Figure 2

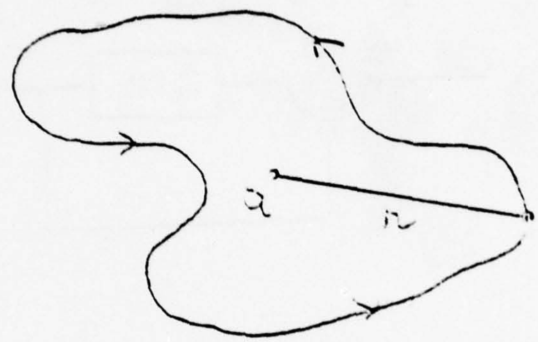
Two closed curves γ_0 and γ_1 are homotopic in $\mathbb{C}$ if there exists a continuous function $\Gamma : I \times I \to \mathbb{C}$ such that:

Intuitively, γ_0 is homotopic to γ_1 if one can continuously deform γ_0 into γ_1. Moreover, it is easily shown that the homotopy relation is an equivalence relation. (4) (5)

Another important property of a closed curve is its index or degree. The index (2) is of closed curve, with respect to a point "a" not in $\{\gamma\}$ is:

$$n(\gamma;a) = \frac{1}{2\pi i} \int_{\gamma} (z-a)^{-1}\, dz$$

This integral measures the net increase in angle that the ray r of Figure 3 accumulates as its tip traverses the trace of γ.



:Figure 3

Intuition for the approach stems in part from the observation that $n(\gamma;-1) = 0$ if and only if γ is homotopic to a point in $\mathbb{C}$- {-1} (cf. prop. 5.4, ref. 2). We will henceforth refer to such a γ as being homotopically trivial. Conversely, γ encircles "-1" if and only if γ cannotbe continuously deformed to a point in $\mathbb{C}$- {-1}. These ideas appear to indicate that the Nyquist encirclement condition is fundamentally a homotopy concept. The tuition is further reenforced when one formulates the Nyquist criterion on the Riemann surface (2) (8) associated with a map, $\hat{f}(s)$. Assuming $\hat{f}(s)$ is analytic on $\mathbb{C}_+$ and bounded at $s=\infty$, the image of simply connected regions in $\mathbb{C}_+$ are simply connected in $\mathbb{C}$. To illustrate the point, let Figure 4-a be the image of the right half plane under $\hat{f}(s)$. The region is not simply connected. Figure 4-b shows the "same region" as it might appear on an appropriate Riemann surface. Here the region is simply connected.

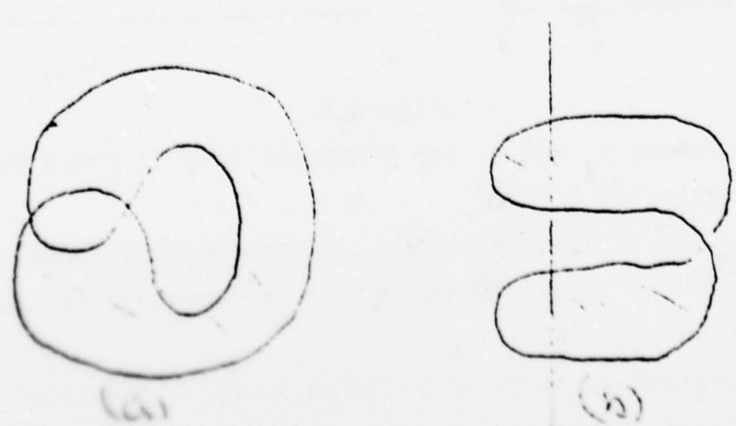


Figure 4

The boundary of the regions depicted in Figure 4 are the Nyquist plots of

$\hat{f}(s)$ in $\mathbb{C}$ and on the Riemann surface. On the Riemann surface the Nyquist test becomes an obvious triviality. In $\mathbb{C}$ it is mathematically more delicate. Our setting uses homotopy theory, a branch of algebraic topology, to establish a topologically invariant relationship between a metric space, X, and an algebraic group called the fundamental group of X, denoted by $\pi(X)$. The relationship is topologically invariant in that homeomorphic spaces have isomorphic fundamental groups.

Specifically, the fundamental group is a set of equivalence classes of closed curves. Each equivalence class consists of a set of curves homotopically equivalent. The group operation is "concatenation" of curves.

For example, the fundamental group of $\mathbb{C}$ consists of one element, $i_{\mathbb{C}}$, the identity, since all closed curves are homotopic to zero. If $X = \mathbb{C} - \{-1\}$, then $\pi(X)$ has a countable number of elements: i_X (the identity) equal to the equivalence class of all closed curves not encircling "-1" and the remaining elements, μ_n (n = 1, 2, 3 ...) consisting of the equivalence class of all closed curves encircling "-1", n times. Moreover, μ_i concatenated with μ_k is equal to the element μ_{k+i}.

Now let X and Y be metric spaces. Let $f: X \to Y$ be locally homeomorphic. In particular, assume that for each point y in Y there exists an open neighborhood G of y such that each connected component of $f^{-1}(G)$ is homeomorphic to G under the map f. Under this condition X is said to be a covering space of Y. (2) (4) Also let $\pi(X)$ and $\pi(Y)$ be the fundamental groups associated with X and Y respectively. With these assumptions, f effects a group isomorphism (i.e. a one to one into mapping preserving group operations) ϕ_f between $\pi(X)$ and a subgroup of $\pi(Y)$ as in the following diagram (4) (5)

$$\begin{tikzcd} X \arrow[r, "F"'] \arrow[d, "f"'] & \Pi(X) \arrow[d, "\phi_f"] \\ Y \arrow[r, "F"'] & N \subset \Pi(Y) \end{tikzcd}$$

Figure 5

F is the functor which establishes the relationship between a topological space and its fundamental group. Finally let us distinguish between a critical point and a critical value. A point z_0 in $\mathbb{C}$ is a critical point of a differentiable function f if $f'(z_0) = 0$. A critical value of f is any point $w = f(z_0)$ whenever z_0 is a critical point.
Now suppose $f:\mathbb{C} \to \mathbb{C}$ is a rational function whose set of poles is $P = \{P_1, \ldots, P_n\}$. Let $Q = \{q_1, \ldots, q_m\}$ be the set of all points in $\mathbb{C}$ such that $f(q_i)$ is a critical value of f. Note that there may be q_i's which are not critical points. To see this consider $g(z) = z^2(z-a)$. $g(0) = 0$ implies "0" is a critical vlue of g, but $g(a) = 0$ with $g'(a) \neq 0$.
Finally, define $T = \{t_i | t_i - f^{-1}(-1), i-1, \ldots, n\}$. Note also that since f is a rational function, P, Q and T are finite sets. Define $X = \mathbb{C} - \{P \cup Q \cup T\}$ and define $Y = f(X)$.
Lemma 1: Under the above hypothesis, X is a covering space of Y. This leads to the following corollary.
Corollary: The fundamental group $\pi(X)$ of X is isomorphic to a subgroup N of $\pi(Y)$.
This corollary says that a closed curve in X is homotopically trivial.

III. THE SCALAR CASE

Let g(s) be as described in the introduction. Appropriately define the sets P, Q, and T and the spaces X and Y so that X is a covering space of Y. Also as per reference (10) and Figure 6, construct the ugly Nyquist contour, χ_R, and the usual Nyquist contour, Γ, where $\Gamma: I \to \mathbb{C} \cup \{\infty\}$

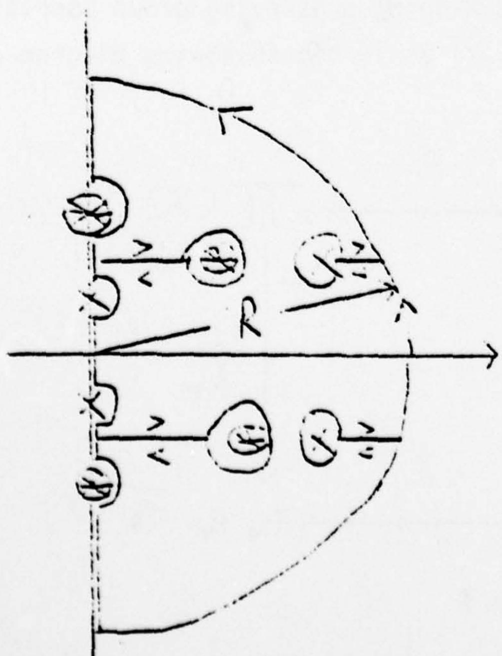


Figure 6-(a)

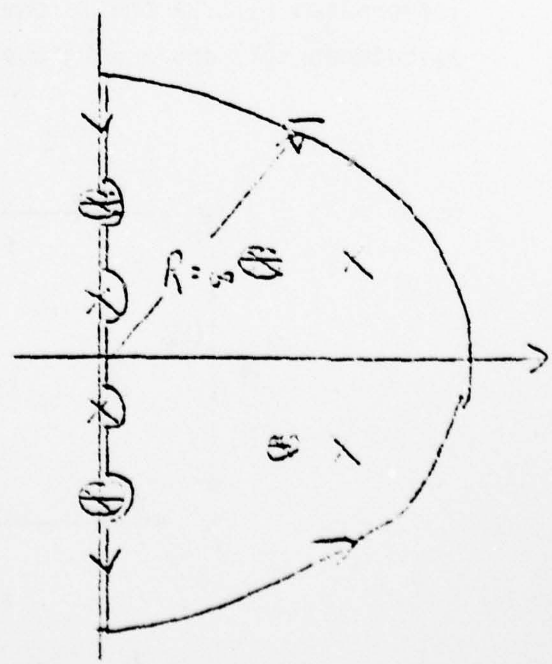


Figure 6-(b)

X indicates a point of P; $\odot$ indicates a point of Q

<u>Lemma 2</u>: Under the above assumptions on $\hat{g}$ and λ_R, $\hat{h}(s)$ is stable if and only if the path $\hat{g}o\lambda_R$ does not encircle "-1". (10)

At this point we must establish this lemma's connection with the classical Nyquist criterion. To this end we compare the information of the Nyquist plot, $\hat{g}o\Gamma$ with the "ugly" Nyquist plot, $\hat{g}o\lambda_R$.

<u>Lemma 3</u>: Let n be the number of poles of $\hat{g}$ in $\mathbb{C}_+$, then

$$\frac{1}{2\pi i}\int_{\hat{g}o\Gamma}(z-1)^{-1}dz = \frac{1}{2\pi i}\int_{\hat{g}o\lambda_R}(z-1)^{-1}dz + n$$

These three lemmas give rise to the following theorem.

<u>Theorem 1</u>: Let $\hat{g}(s)$ be as above. Then $\hat{h}(s)$ is stable if and only if the Nyquist plot of $\hat{g}(s)$ does not pass through "-1" and encircles "-1" exactly n times where n is the number of poles of g(s) in $\mathbb{C}_+$.

IV. MATRIX CASE

Let the entries of an nxn matrix $\hat{G}(s)$ be rational functions in the complex variable s. Suppose $\hat{G}(s)$ characterizes the open loop gain of the single loop feedback system of Figure 7.

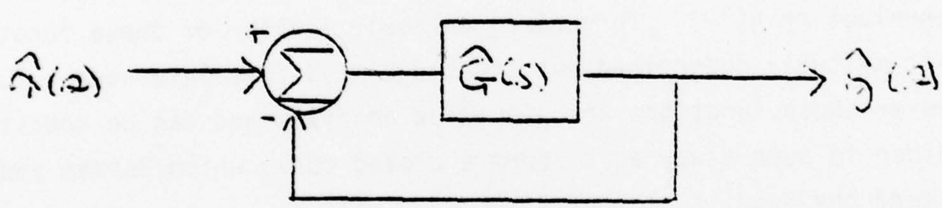


Figure 7

$\hat{x}(s)$ and $\hat{y}(s)$ are n vectors whose entries are also rational functions of s which represent the input and output of the system respectively.

This article assumes each entry of $\hat{G}(s)$ is bounded at $X = \infty$. Thus $\hat{G}(s)$ as a mapping, $\hat{G}(\cdot):\mathbb{C} \to \mathbb{C}^{nxn}$, is analytic on $\mathbb{C}$ except at a finite number of points, the poles of its entries.

For Figure 7 to be well defined we require that det $[I+\hat{G}(s)]=0$ Thus there exists a closed loop convolution operator, H, such that $y = H*x$. Moreover the Laplace transform of H, $\hat{H}(s)$ satisfied

$$\hat{H}(s) = \hat{G}(s)[I+\hat{G}(s)]^{-1}$$

For this system to be stable, $\hat{H}(s)$ must have all its poles in $\mathbb{C}_-$ and have all its entries bounded at $s = \infty$.

Under the assumptions on $\hat{G}(s)$, the following factorization is valid:

$$\hat{G}(s) = N(s)D^{-1}(s)$$

where $N(s)$ and $D(s)$ are right co-prime, polynomial matrices in s with $\det[D(s)] \neq 0$. Moreover s_0 is a pole of $\hat{G}(s)$ if and only if it is a zero of $\det[D(s)]$. (9)

Dosoer and Schulman (3) have shown that the close loop operator H is stable if and only if $\det[N(s)+D(s)] \neq 0$ for s in $\mathbb{C}_+$ and $\det[I+G(\infty)] \neq 0$.

Using this fact, we state and prove the following:

Theorem 2: H is stable if and only if (1) the Nyquist plot of $\det[N(s)+D(s)]$ does not encircle nor pass through "0", and (2) $\det[I+\hat{G}(\ )] \neq 0$. (10)

Observe that if one assumes the open loop gain to be stable (i.e. $\hat{G}(s)$ has all poles in $\mathbb{C}_+$) then $\det[I+G(s)]$ in the above theorem. This follows since for all s in $\mathbb{C}_+$, $\det[N(s)+D(s)] = \det[I+\hat{G}(s)] \det[D(s)]$ with $\det[D(s)] \neq 0$. Thus in $\mathbb{C}_+$ $\det[N(s)+D(s)]$ has a zero if and only if $\det[I+\hat{G}(s)]$ has a zero.

Finally, it is worthwhile to point out the relationship between the above formulated multivariable Nyquist criterion and that formulated by Barman and Katznelson. For this purpose we let $\lambda_j(iw)$; $j=1, \ldots, n$; denote the n eigenvalues of $\hat{g}(iw)$. In general parameterization of these function by iw is not uniquely determined but one can always formulate such a function. Moreover these functions are piecewise analytic and can be concatonated together in such a way as to form a closed curve which Barman and Katzelson term the Nyquist plot of $\hat{G}(s)$.

Now, since

$$\det[I + \hat{G}(iw)] = \prod_{j=1}^{n} [1 + \lambda_j(iw)]$$

and the degrees of a product is the sum of the degrees of the individual factors and also equals the degree of the cancatonation of the factors, the degree of the Barman and Katznelson plot with respect to "-1" coincides with the degree of our plot with respect to "0". As such, even though the two plots are different their degrees coincide and hence either can be used for a stability test.

Acknowledgement: The authors would like to acknowledge the contribution of Dr. John Murray (Dept. of Mathematics, Texas Tech University) whose continuous flow of counter examples shaped the ideas presented herein.

REFERENCES

1. Barman, John F. and Katzenelson, Jacob, "A Generalized Nyquist-type Stability Criterion for Multivariable Feedback Systems," Int. Journal of Control, 1974, Vol. 20, pp. 593-622.

2. Conway, John B., "Functions of One Complex Variables," Springer-Verlag, New York, 1974.

3. Desoer, Charles A., and Schulman, J. D., "Cancellations in Multivariable Continuous-time and Discrete-time Feedback Systems," Memorandum No. ERL-M346, Berkeley, College of Engrg., Univ. of CA, 1972.

4. Hocking, John G. and Young, Gail S., "Topology" Addison-Wesley Inc., Reading, Mass, 1961.

5. Massey, William S., "Algebraic Topology: An Introduction, "Harcourt, Brace and World Inc., New York, 1967.

6. Milnor, John W., "Topology from the Differential Viewpoint," Univ. Press of Virginia, Charlottesville, 1965.

7. Rudin, Walter, "Functional Analysis," McGraw-Hill, Inc., New York, 1973.

8. Springer, George, "Introduction to Riemann Surfaces," Addison-Wesley Inc., Reading, Mass., 1957.

9. Wang, S. H., "Design of Linear Multivariable Systems," Memorandum No. ERL-M309, Electronics Research Laboratory (Berkeley, CA, University of CA) 1971.

10. DeCarlo, R. and Saeks, R., The Encirclement Condition: An Approach Using Albegraic Topology, to be published.

THE "FOURIER" TRANSFORM OF A RESOLUTION SPACE AND A THEOREM OF MASANI*

R.A. DeCarlo, R. Saeks and M.J. Strauss

*This research supported in part by AFOSR grant 74-2631.

THE "FOURIER" TRANSFORM OF A RESOLUTION SPACE AND A THEOREM OF MASANI*

R. A. DeCarlo, R. Saeks, and M. Strauss
Texas Tech University
Lubbock, Texas

ABSTRACT

Using two classic theorems (one of Mackey and another of Strone) and a recent result of Masani and Rosenberg, this paper pieces together a generalized frequency response theory for an abstract Uniform Resolution Space. The present theory assimilates past work as done by Falb, Freedman, Anton, Masani and Rosenberg, and one of the authors. The results of the paper are not new, but are merely a rearrangement of subtleties uncovered by the aforementioned authors. An interesting consequence of this work was that an abstract Uniform Resolution Space has both a "time transform" and a "frequency transform". Such a duality is not readily identifiable in an L_2 function space since the time transform, there, is the identity.

INTRODUCTION

Fourier analysis is basic to the design and understanding of physical systems. The property that convolution in the time domain maps into a product in the frequency domain, yields a theory both practical and aesthetically pleasing. This note provides what is hoped to be a generalized frequency response theory for arbitrary, closed, linear, time invariant operators on a uniform resolution space. Previous attempts at providing a general frequency theory have illuminated numerous subtleties, yet still appear inadequate for one reason or another. Interestingly enough, the mathematics necessary for such a synthesis is well entrenched in the literature. This paper merely pieces these results together and reinterprets them in light of the work of Falb, Freedman, Masani, Rosenberg and Saeks.[3,9,12,21]

This research supported in part by Air Force Office of Scientific Research Grant AFOSR 74-2631.

Classical Fourier analysis consists essentially of two fundamental ideas--the idea of a "transform" from time to frequency and the property of a time-invariant mapping to a product of functions in frequency. We desire a Fourier representation for time invariant operators defined on an appropriate space. Two avenues arise. A traditional approach uses a Fourier-like integral to obtain the representation. In an abstract approach, the Fourier representation is a spectral representation of the abstract operator relative to an appropriate spectral measure. This road is both more general and eliminates the need for a specific representation of the operator.

Falb, Freedman and Anton[3,5] developed a generalization closely paralleling the classical theory. The formulation considers Hilbert space-valued L_2 functions (square integrable relative to the Haar measure), defined over a locally compact abelian (LCA) group, G, and operators which are characterized by an L_1 convolutional weighting function. The theory is highly representation-dependent and fits awkwardly into the setting of an abstract resolution space. In fact, the identity and unit delay are not admissable to the theory. The major advantage is that one obtains an operator-valued Fourier representation.

Masani and Rosenberg[11,7] use a spectral theoretic vehicle to alleviate the difficulty of a specific representation of the operator. Moreover, the theory settles nicely into an abstract setting. Yet, the frequency response is always scalar-valued, even in the multivariable case, and the concept of a "transform" is absent.

Finally, Saeks[21] has a Masani-like development whose Fourier representation assumes values in a suitably restricted class of operators. The advantages are the compatibility with abstract spaces and an operator-valued frequency response. Yet still, the concept of a transform is missing and major existence questions are still present.

The structure of the present theory rests on the classic theorems of Mackey[7] and Stone[4] and a recent theorem of Rosenberg and Masani[10]. With this comment, we define the setting.

UNIFORM RESOLUTION SPACE

A resolution space is a pair, (H,E) where H is a Hilbert space and E is a spectral measure on an ordered LCA topological group, G. On an ordered LCA group, a spectral measure determines a resolution of the identity, and conversely. Thus, it is advantageous to work with the resoltuion of the identity $E^t = E([-\infty,t])$, rather than with the spectral measure E, as illustrated at the end of this section.

As an example, consider the Hilbert space, L_2, together with the truncation operator, E^t, defined as

$$(E^t x)(q) = \begin{cases} x(q) & q \leq t \\ 0 & q > t \end{cases}$$

or equivalently, the spectral measure, defined via

$$(E(B)x)(q) = \begin{cases} x(q) & q \in B \\ 0 & q \notin B \end{cases}$$

for all Borel sets B.

In addition L_2 admits a group U of shift operators U^t, defined as

$$(U^t x)(q) = x(q - t).$$

Thus, the concept of time invariance is well defined in a classical L_2 setting. In general, a resolution space lacks the concept of time invariance. Such a property requires an extension of the concept of the L_2 "time-shift". A group of such operators, in general, fails to exist in an arbitrary resolution space.

In particular, we seek a strongly continuous group of unitary operators (i.e., $U^{t-s} = U^t(U^s)^{-1}$ for all t and s in G), such that

$$U^t E(B) = E(B + t)U^t$$

for all t in G and Borel sets B. A resoltuion space, together with such a group U of shift operators, U^t, is a Uniform Resolution Space (URS), denoted by the triple (H,E,U).

Underlying each URS is an ordered LCA group, G, which, for our purposes, is time. Associated with G is a "character group", $\hat{G}$, which is the group of continuous homomorphisms from G into the multiplicative group of complex numbers of magnitude one. Note that $\hat{G}$ is, in general, not ordered.

In like manner, attached to each URS (e.g., (H,E,U)), defined over G, is a "dual" character space $(H,\hat{U},\hat{E})$*, defined over $\hat{G}$. $\hat{E}$ and $\hat{U}$ are a spectral measure and a group of shift operators, respectively, defined via the two equalities

$$U^t = \int_{\hat{G}} (\gamma,-t)\,d\hat{E}(\gamma) \qquad t \in G$$

and

$$\hat{U}^{\gamma} = \int_{\hat{G}} (\gamma,-t)\,dE(t) \qquad \gamma \in \hat{G}.$$

Here, $(\gamma,-t)$ denotes the complex number of magnitude one, resulting from the operation of the character γ in $\hat{G}$ acting on -t in G, and where the integral is the Lebesque integral. Stone's theorem(4) assures the existence and uniqueness of $\hat{E}$ and $\hat{U}$.

Oddly, the character space $(H,\hat{U},\hat{E})$ is not a resoltuion space since $\hat{G}$ is not ordered. However, $(H,\hat{U},\hat{E})$ displays all the resolution space properties which do not depend on the ordering of G. In fact, by Stone's theorem(6),(7),(12), $\hat{U}$ is a group of shift operators for $\hat{E}$, satisfying the imprimitivity equality over G--i.e.,

$$\hat{U}^{\gamma}\hat{E}(B) = \hat{E}(B+\gamma)\hat{U}^{\gamma}.$$

*We have adopted the ordering $(H,\hat{U},\hat{E})$ because via Stone's Theorem, E and $\hat{U}$ contain the same information. Moreover, U and $\hat{E}$ do. Thus, $(H,\hat{U},\hat{E})$ rather than $(H,\hat{E},\hat{U})$.

For our purposes, the character group plays the role of frequency.

Now, the physical properties of causality, memorylessness, time invariance, etcetra, have precise descriptions in the uniform resolution space structure. In particular, for bounded operators, T, on (H,E,U), causality is equivalent to $E^tT = E^tTE^t$; [1,2,20]; anticausality, to $E_tT = E_tTE_t$*; memorylessness, to $E^tT = TE^t$ which, in turn, is equivalent to T, being both causal and anticausal. Since memorylessness is a symmetric concept, it has an analog in the character space, $(H,\hat{U},\hat{E})$, whereas causality does not. Because of this, we say a bounded operator, T, is time invariant if $\hat{E}(B)T = T\hat{E}(B)$ for all Borel sets B of $\hat{G}$. Via Stone's theorem, this is equivalent to $U^tT = TU^t$ for all t in G. Clearly, we emphasize the character space in the definition of time invariance.

In the case of unbounded operators, (e.g., the derivative operator), T is causal if $E^tT \subseteq E^tTE^t$**; T is anticausal if E_tT _ E_tTE_t; T is memoryless if $E^tT \subseteq TE^t$; and, finally, although somewhat non-intuitively, T is time invariant if $\hat{E}(B)T \subseteq T\hat{E}(B)$ for all Borel sets B in $\hat{G}$, where, again, we emphasize the definition in the character space. For unbounded operators, Stone's theorem, in general, does not yeild an equivalent statement (such as $U^tT = TU^t$) in the original resoltuion space. However, for the case of linear, single-valued, closed operators with domain dense in H, then $U^tT = TU^t$ if and only if $E^tT \subseteq TE^t$.[9,10] The fundamental role of the character space becomes more clear in the following section.

EQUIVALENT SPACES

In this section, Mackey's theorem verifies an equivalence between an abstract URS, (H,E,U) and a function space, $(L_2(G,K),\ X_B, \sigma^t)$. Now, the relevant information contained in (H,E,U) is also contained in $(H,\hat{U},\hat{E})$. Thus, applying Mackey's theorem to (H,E,U) (under the guise of $(H,\hat{U},\hat{E})$), another equivalence to $(L_2(\hat{G},K),\ \sigma^\gamma,\ X_{\hat{B}})$ exists. Furthermore, $(L_2(G,K),\ X_B,\ \sigma^t)$ and $(L_2(\hat{G},K),\ \sigma\gamma,\ X_{\hat{B}})$ have and affinity via Stone's theorem.

---

*$E_t = E((t,\infty]) = I - E^t$.

**For an unbounded operator, T, on a resolution space, (H,E,U), the domain of E^tT is smaller than the domain of TE^t. As such, the containments indicate that, where the domains of E^tT and TE^t coincide, then $E^tT = TE^t$.

After numerous references throughout this development to the above authors, we, at last, precisely state their results. Hopefully, this will facilitate understanding of the maps between the various spaces, hinted to in the above paragraph. The following is a statement of Stone's theorem for LCA groups.[4,14]

Suppose G is an LCA group and $\hat{G}$, its "dual" character group; let $(\gamma, -t)$ be the complex number of magnitude one, resulting from the operation of γ in G on $-t$ in G; define $\hat{\Sigma}(\Sigma)$ as the σ-algebra of Borel sets of $\hat{G}(G)$; finally, let $\{U^t | t \text{ in } G\}$ $(\{\hat{U}^\gamma | \gamma \text{ in } \hat{G}\})$ be a strongly continuous group of unitary operators on a complex Hilbert space, H, onto H. Then, there exists a unique spectral measure, $\hat{E}(\cdot)$ $(E(\cdot))$, for H on $\hat{\Sigma}(\Sigma)$, such that all t in G (γ in $\hat{G}$),

$$U^t = \int_{\hat{G}} (\gamma, -t)\, d\hat{E}(\gamma)$$

or

$$\hat{U}^\gamma = \int_{G} (\gamma, -t)\, dE(t).$$

The initial task, now is to construct an equivalence between two L_2 spaces via this theorem.

Consider the URS, $(L_2(G,K), X_B, \sigma^t)$, and the character space, $(L_2(\hat{G},K), \sigma^\gamma, X_{\hat{B}})$, where G is an LCA group; $\hat{G}$, its character group, X_B, the characteristic function of the Borel set B in Σ; σ^t, the classical shift operator (i.e., $(\sigma^t f)(q) = f(q - t)$*, and, lastly, the measure on the space will be the Haar measure, m.

Now, the Fourier transform maps $L_2(G,K)$ to $L_2(\hat{G},K)$ in such a manner that X_B is taken to the spectral measure on $L_2(\hat{G},K)$, whose integral is σ^γ. Similarly, σ^t maps to the unitary group on $L_2(\hat{G},K)$, whose associated spectral measure is $X_{\hat{B}}$. As such, $(L_2(\hat{G},K), \sigma^\gamma, X_{\hat{B}})$ is the Fourier transform of the character space for $(L_2(G,K), X_B, \sigma^t)$.

---

*Here, $\{\sigma^t | t \text{ in } G\}$ and $\{X_B | B \text{ in the set of Borel sets of } G\}$ serve as the strongly continuous group of shift operators and the spectral measure, respectively.

Interpreting this, we have X_B completely determining σ^γ and $X_{\hat{B}}$, completely specifying σ^t. The link between these two spaces and an abstract resolution space is Mackey's theorem. The statement of his theorem for LCA groups follows:[19]

Let E be a spectral measure on the Borel sets of an LCA group, J, and let U be a strongly continuous, unitary representation of J, such that

$$E(B + t)U^t = U^t E(B)^{**}$$

for all Borel sets B of J and all t in J; then, there exists a unique Hilbert space, K, and a unitary transformation, M,

$$M: H \to L_2(J,K,m)$$

such that

(1) $ME(B)M^{-1} = X_B$ for all Borel sets of J; and

(2) $MU^tM^{-1} = \sigma^t$ for all t in J,

where K is a complex Hilbert space and m, the Haar measure.

For an arbitrary URS, (H,E,U), defined over an LCA group, G, by design, E and U satisfy the imprimitivity equality. Moreover, the character space, $(H,\hat{U},\hat{E})$, possess the property that $\hat{E}$ and $\hat{U}$ satisfy the imprimitivity equality.[20] Finally, it is clear that $(L_2(G,K),\ X_B,\ \sigma^t)$ and $(L_2(\hat{G},K),\ \sigma^\gamma,\ X_{\hat{B}})$ satisfy the hypothesis of the theorem. Therefore, by blending Mackey's and Stone's theorem, the following commutative diagram results:

$$\begin{tikzcd}
 & (H,\ E(\hat{U}),\ U(\hat{E})) \arrow[dr, "\hat{M}"] & \\
(L_2(G,K),\ X_B,\ \sigma^t) \arrow[ur, "M"] \arrow[rr, leftrightarrow, "F"] & & (L_2(\hat{G},K),\ \sigma^\gamma,\ X_{\hat{B}}).
\end{tikzcd}$$

Figure 1

Remarkably, the diagram reveals that an arbitrary URS is equivalent, under a memoryless, time invariant, unitary transformation--a uniform resolution space insomorphism, to an L_2 space. Distinctions among the spaces for a fixed G,

---

**This is the imprimitivity equality.

AD-A055 613 TEXAS TECH UNIV LUBBOCK INST FOR ELECTRONICS SCIENCE F/G 12/1
RESOLUTION SPACE, NETWORKS, AND NON-SELF-ADJOINT SPECTRAL THEOR--ETC(U)
1978 R SAEKS, L TUNG, R A DECARLO AFOSR-74-2631
UNCLASSIFIED AFOSR-TR-78-1045 NL

therefore, depend only on the cardinality of the space, K.[19] Mixing this equivalence with a recent result of Masani and Rosenberg[10], gives the desired structure--i.e., "time invariance" maps to multiplication.

APPLICATION OF THE MASANI-ROSENBERG RESULT

This section begins with the result of the above mentioned authors. The theorem is not stated in its general form[10], but is restricted to a group J, a Hilbert space, K, and the Haar measure, m.

Let T be a closed, single-valued, linear operator with dense domain on $L_2(J,K,m)$, such that T commutes with the operation of multiplication by the characteristic function of a Borel set--i.e.,

$$\chi_B T \subseteq T\chi_B, \quad \text{for all } B \text{ in } \Sigma \text{ (the } \sigma\text{-algebra of all Borel sets of } J\text{)}.$$

The, there exists a measurable function, T on J, whose values are operators on K, such that

$$(Tf)(j) = T(j)f(j) \qquad j \text{ in } J.$$

This theorem applies to function space. Thus, to verify the sought after properties on the abstract URS, we fist apply the Mackey Transforms as in Figure 1, redrawn below for simplicity.

$$\begin{tikzcd}[column sep=small] & (H, E(\hat{U}), U(\hat{E})) \arrow[dl, leftrightarrow, "M"'] \arrow[dr, leftrightarrow, "\hat{M}"] & \\ (L_2(G,K), \chi_B, \sigma^t) \arrow[rr, leftrightarrow, "F"] & & (L_2(\hat{G},K), \sigma^\gamma, \chi_{\hat{B}}). \end{tikzcd}$$

Our discussion dwells upon two types of operators in the abstract URS, memoryless operators and time invariant. First, we consider the time invariant case.

Let T be a closed, linear, single-valued, time invariant operator on $(H,E(U), U(\hat{E}))$. Recall that T is time invariant if $\hat{E}(\hat{B})T \subseteq T\hat{E}(\hat{B})$. Under the

Mackey Transform, $\hat{M}$ (which we term the Mackey frequency transform), we have an equivlent statement in $(L_2(\hat{G},K), \sigma^\gamma, X_B)$, as follows:

$$X_B T_{\hat{M}} \subseteq T_{\hat{M}} X_B,$$

where $T_{\hat{M}}$ is the image of T under the $\hat{M}$ transformation. Clearly, the conditions of the Masani-Rosenberg theorem are satisfied. Thus, there exists a mapping, $\hat{T}:\hat{G}\to K$, such that

$$(Th)(\gamma) = \hat{T}(\gamma)h(\gamma)$$

for all γ in $\hat{G}$ and h in $L_2(\hat{G},K)$. This says that time invariant, closed, linear, single-valued operators on an abstract URS are, as hoped, multiplications in the "frequency domain"--i.e., in $(L_2,(\hat{G},K),\sigma^\gamma, X_{\hat{B}})$.

Now, let T be a memoryless, linear, closed, single-valued operator on (H,E,U). Recall that T is memoryless if $E^tT = TE^t$. Thus, by reasoning similar to the time invariant case, the image of T under the Mackey-time transform, M, commutes with X_B in $(L_2(G,K), X_B, \sigma^t)$. Thus, it is equivalent to a multiplication by the Masani-Rosenberg theorem.

This structure, then, shows that certain operators on an abstract URS have the "right" properties. It is interesting to note that there is a duality inherent in this formulation. The presence of a Mackey "time-transform" and a corresponding "frequency-transform" is apparently necessary for the cohesiveness of the theory.

CONCLUSIONS

The above ideas assimilate past theories in a number of ways. The theory generalizes the Falb-Freedman-Anton work because of the abstract setting and because it is valid for a larger class of operators. Clearly, there is no restriction to scalar-valued frequency responses as in (9) and (13). Moreover, it circumvents the existence questions associated with Saeks' work.[21] In fact, as in (21), given appropriate conditions, a multiplication on a function space can be viewed as an integral over the spectral measure, defined via multiplication by X, i.e.

$$T = \int T(\omega)d\chi(\omega).$$

Hence, the pre-image of T under the Mackey frequency transform in (H,E,U) takes the form

$$T = \int \hat{T}(\gamma)d\hat{\Xi}(\gamma),$$

as was specified in (21).

In a private conversation with one of the authors, Desoer raised a question about the fact that differential operators satisfied the definition of memorylessness in an abstract setting. The question caused some doubt in our minds as to the appropriateness of the definition. This note gives a parital answer, in that closed, memoryless operators are multiplications. Hence, the apparent pathology noted by Desoer can only arise in the case of nonclosed operators.

REFERENCES

1. R. DeSantis, Causality and Stabilty in Resolution Space, Proc. 14th Midwest Symp. on Circuit Theory, Univ. of Denver, 1971.

2. __________, Causality Structure of Engineering Systems, Doctoral Thesis, Univ. of Michigan, Ann Arbor, 1971.

3. P.L. Falb and M.I. Freedman, A Generalized Transform Theory for Causal Operators, SIAM J. Control, 7 (1969) pp. 452-471.

4. P.A. Filmore, Notes on Operator Theory, Van Nostrand-Reinhold, New York, 1970.

5. M.I. Freedman, P.L. Falb and J. Anton, A Note on Causality and Analyticity, SIAM J. Control 7 (1969), pp. 472-478.

6. G.W. Mackey, Induced Representation of Groups and Quantum Mechanics, W.A. Benjamin, New York, 1968.

7. __________, A Theorem of Stone and von Neumann, Duke Math. J., 16 (1949), pp. 313-325.

8. __________, Group Representations, Lecture Notes, Oxford Univ., 1966-1967.

9. P. Masani and M. Rosenberg, When is an Operator the Integral of a Spectral Measure, (to appear).

10. _______________, The Multiplication Operator in L_2 over a Localizable Space and Bochner's Theroem, (to appear).

11. P. Masani, The Normality of Time-Invariant Subordinate Operators in a Hilbert Space, Bull. Amer. Math Soc., 71 (1965), pp. 546-550.

12. Orthogonally Scattered Measures, Advances in Math., 2 (1968), pp. 61-117.

13. ________, Quasi-Isometric Measures and their Applications, Bull. Amer. Math. Soc. 76 (1970), pp. 472-528.

14. F. Riesz and B. Sz-Nagy, Functional Analysis, Fredrick Ungar, New York, 1963.

15. R. Saeks, Causality in Hilbert Space, SIAM Review, 12 (1973), pp. 283-308.

16. ________, State in Hilbert Space, SIAM Review, 16 (1973), pp. 283-308.

17. ________, On the Existence of Shift Operators, Tech. Memo EE-707, Univ. of Notre Dame, Indiana, 1970.

18. ________, Causal Factorization, Shift Operators and the Spectral Multiplicity Function, Vector and Operator Valued Measures and Applications, Academic Press, Inc. 1973.

19. ________, Resolution Space Operators and Systems, Springer-Verlay, New York, 1973.

20. ________, Fourier Analysis in Hilbert Space, SIAM Review, 15 (1973), pp. 604-637.

21. Segal, I.E., Equivalence of Measure Spaces, Amer. J. Math, 73 (1951), pp. 275-313.

AN APPROACH TO RESOLUTION SPACE USING RELATIVISTIC TIME*

D. Brandon

*This research supported in part by AFOSR grant 74-2631.

CHAPTER I

INTRODUCTION

One technique which has been developed in recent years for the operator theoretic study of systems is the use of resolution space [1]. The basic motivation behind the development of resolution space was to overcome the impossibility of defining time-based concepts such as causality in the Hilbert and Banach spaces which are the setting for classical operator theory.

Resolution space techniques have been very successful in achieving the goal of including time-based concepts in operator theoretic systems theory, but the concept of time which has been used is classical in nature. It has been known since the early part of this century that an accurate model of the physical universe must be based on the concept that space and time are intimately connected. This is the central thesis of the theory of relativity formulated by Einstein [2]. Classical resolution space techniques ignore this connection between space and time, and thus, it might be suspected that an extension of the resolution space concept to include the constraints posed by the theory of relativity could offer new insights.

Due to the difficulty of merging the operator theory required by resolution space and the theory of differentiable manifolds required by general relativity, the relativistic resolution space theory is developed only for the case comparable to the classical resolution space development based on Hilbert spaces such as $H = L_2(G,K,\mu)$, the Hilbert space of functions defined on an ordered, locally compact, abelian group G which take values in a Hilbert space K, and which are square integrable relative to a Borel measure μ. In H, we can define a spectral measure E by multiplication by the characteristic function

1.1. $[E(A)f](s) = \chi_A(s)f(s)$

for each Borel set A. Given the spectral measure E, we can define a resolution of the identity by

1.2. $E^t = E(-\infty,t)$, $t \in G$.

In this case, E^t reduces to a family of truncation operators

1.3. $(E^t f)(s) = \begin{cases} f(s) & ; s \leq t \\ 0 & ; s > t. \end{cases}$

We also have a resolution of the identity E_t defined by

1.4. $E_t = I - E^t$.

The pair (H,E) is the classical L_2 resolution space in which the resolution of the identity E^t allows the introduction of the desired time-based concepts.

We define space-time to be a pair (M,g) where M is a real, four-dimensional, connected C^{∞} Hausdorff manifold, and g is a globally defined C^{∞} tensor field of type $(0,2)$ which is nondegenerate and Lorentzian. Then analogously to the above definition for classical L_2 resolution space, we can define relativistic L_2 resolution space to be the pair (H,E) where $H = L_2(M,K,\mu)$ is the Hilbert space of functions defined on M with values in a Hilbert space K, and which are square integrable with respect to a Borel measure μ, and E is a spectral measure defined by multiplication by the characteristic function

1.5. $$[E(A)f](s) = \chi_A(s)f(s)$$

for each Borel set A.

In order to proceed further, it is necessary to find some counterpart to the resolution of the identity induced in the classical case by the spectral measure E. We don't have a resolution of the identity in the relativistic case since the manifold M isn't ordered. It is still possible to parallel the classical resolution space development if we first define the past and future of any point x in M. Although some care must be taken to obtain a precise definition, the past of a point x is essentially the set of all points in M which could have sent a signal in the past traveling at a speed less than or equal to the speed of light which could be received by x in the present. Similarly, the future of x is the set of all points in M which could receive a signal sent by x traveling at a speed less than the speed of light. Then the family of projections E^x is defined by

1.6. $$E^x = E[J^-(x)]$$

where $J^-(x)$ is the past of x, and similarly, the family of projections E_x is defined by

1.7. $$E_x = E[J^+(x)]$$

where $J^+(x)$ is the future of x.

In addition to the fact that E^x and E_x don't form resolutions of the identity, we also have

1.8. $$E^x + E_x \neq I \text{ (Since } J^+(x) \cup J^-(x) \neq M\text{)},$$

whereas in the classical case

1.9. $$E^t + E_t = I \ .$$

The lack of an order on M, and the noncomplementary nature of E^x and E_x combine to make the relativistic definitions of concepts such as causality and strict causality more complicated, and definitions which were equivalent in the classical case are no longer equivalent in the relativistic case. It also turns out that strictly causal, strictly anticausal, and memoryless operators are no longer enough for decomposition of an arbitrary operator, and the totally new concept of a spacelike operator must be introduced. The increased complexity of the relativistic case also prevents several of the classical theorems from carrying over into the relativistic setting.

The most familiar space-time is the Minkowski space-time of special relativity ([2],[3],[4]). Minkowski space-time is the manifold R^4 with a flat Lorentz metric g. If $x = (x_1,x_2,x_3,x_4)$ and $y = (y_1,y_2,y_3y_4)$ are two points in Minkowski space-time, then

1.10. $$g(x,y) = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4.$$

A nonzero point x is said to be <u>timelike</u> of $g(x,x)<0$, <u>spacelike</u> if $g(x,x)>0$, and <u>null</u> if $g(x,x) = 0$. For a given point x, all the points separated from x by a timelike or null distance form a hypercone called a <u>lightcone</u>. If we assume that all timelike distance vectors can be classified as either future-directed or past-directed, then the lightcone can be divided into two parts. Points separted from x by a future-directed timelike distance are said to be in the <u>future</u> of x and points separated from x by a past-directed nonspacelike distance are said to be in the <u>past</u> of

x. All other points are said to lie in the spacelike region about x. With two spatial dimensions supressed, the light-cone of a point x is shown in Figure 1.

In the next chapter, we develop a relativistic resolution space theory for the special case in which the space-time manifold is a two-dimensional version of Minkowski space-time. We first define a special set of lines called null lines. These lines are essentially the paths along which a light ray would travel in our space-time. Then the past and future of a null line are defined, and these concepts are used to define a special class of sets called diamond sets. The diamond sets are shown to form a semi-ring and this permits them to be used to establish an integration theory paralleling that used in the classical resolution space development ([1],[5]).

Next, causal, anticausal, and memoryless operators are defined, and then shown to have properties similar to those in the classical development [1].

After this, strictly causal and strictly anticausal operators are defined, and once again, the development in the classical case is paralleled [1].

Finally, the decomposition theorems for an arbitrary operator are stated and proved [1]. In order to obtain a complete decomposition, a new type of operator called spacelike is introduced. This operator essentially takes

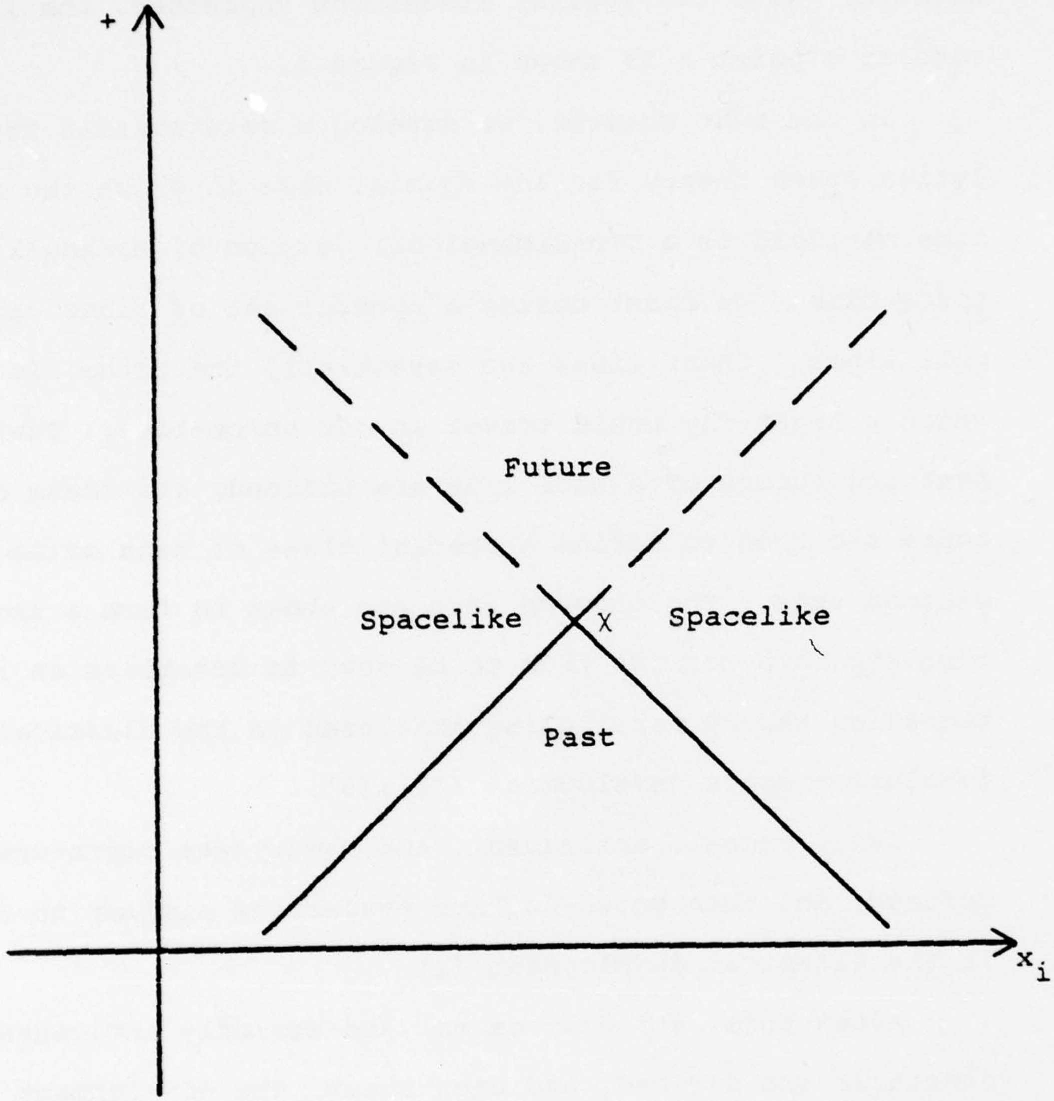


Figure 1.

care of the spacelike region ignored by the causal and anticausal operators.

The conclusion discusses the problems involved in extending the two-dimensional theory to a more general case. A possible technique for developing a general theory is indicated, and we state a conjectured property of the spacelike operators.

CHAPTER II

THE TWO DIMENSIONAL CASE

Our space-time will be an ordered pair (M,g) where M is the connected, two-dimensional (one space and one time coordinate), Hausdorff C^{∞} manifold $R_1{}^2$ and g is the usual flat Lorentz metric used in special relativity, i.e., for (x_1,t_1), $(x_2,t_2) \varepsilon M$,

2.1. $$g[(x_1,t_1),(x_2,t_t)] = x_1x_2 - t_1t_2.$$

This space is essentially R^2 with an "inner product" defined by g. Following Penrose [6], a non-zero tangent vector X is said to be <u>timelike</u> if $g(X,X)<0$, <u>spacelike</u> if $g(X,X)>0$, and <u>null</u> if $g(X,X) = 0$. A C^1 curve in M is called <u>timelike</u> (<u>spacelike</u> , <u>null</u>) if the tangent vector to the curve at each point is timelike (spacelike, null). A curve will be called <u>non-spacelike</u> if the tangent vector at each point of the curve is either timelike or null.

We now assume that we can divide the non-spacelike vectors at each point in M into two groups which we will call future- and past- directed non-spacelike vectors. Essentially, a non-spacelike vector is future-directed if it makes an angle of 45° to 135° with the x-axis, and it is past-directed if it makes an angle of 225° to 315° with the x-axis.

Let $x = (x,t)$ be a point of M. The future of x will be denoted $J^+(x)$ and is defined to be the set of all points of M which can be reached from x by a future-directed timelike curve, i.e., a curve whose tangent vectors are all future-directed timelike vectors. $J^+(x)$ doesn't include x. The past of x will be denoted $J^-(x)$ and is the set of all points of M which can be reached from x by a past-directed non-spacelike curve, i.e., a curve in M whose tangent vectors are all past-directed non-spacelike vectors. $J^-(x)$ does include x. For an illustration, see Figure 2.

It is easily seen that the boundaries of the past and future of a point are lines with slopes of ± 1. Since these lines are so important, they will be given a name, null lines. The future of a null line is defined to be the set of all points of M which lie in the future of some point on the null line. The past of a null line is defined to be the set of all points of M which lie in the past of some point on the null line. See Figure 3.

In order to carry through the integration theory in our space-time setting, we need to define the class of sets which will be used to partition M. A diamond set will be defined to be the intersection of the pasts or futures of any finite set of null lines, the empty set, or the whole space M. The future (past) of a diamond set is the union of

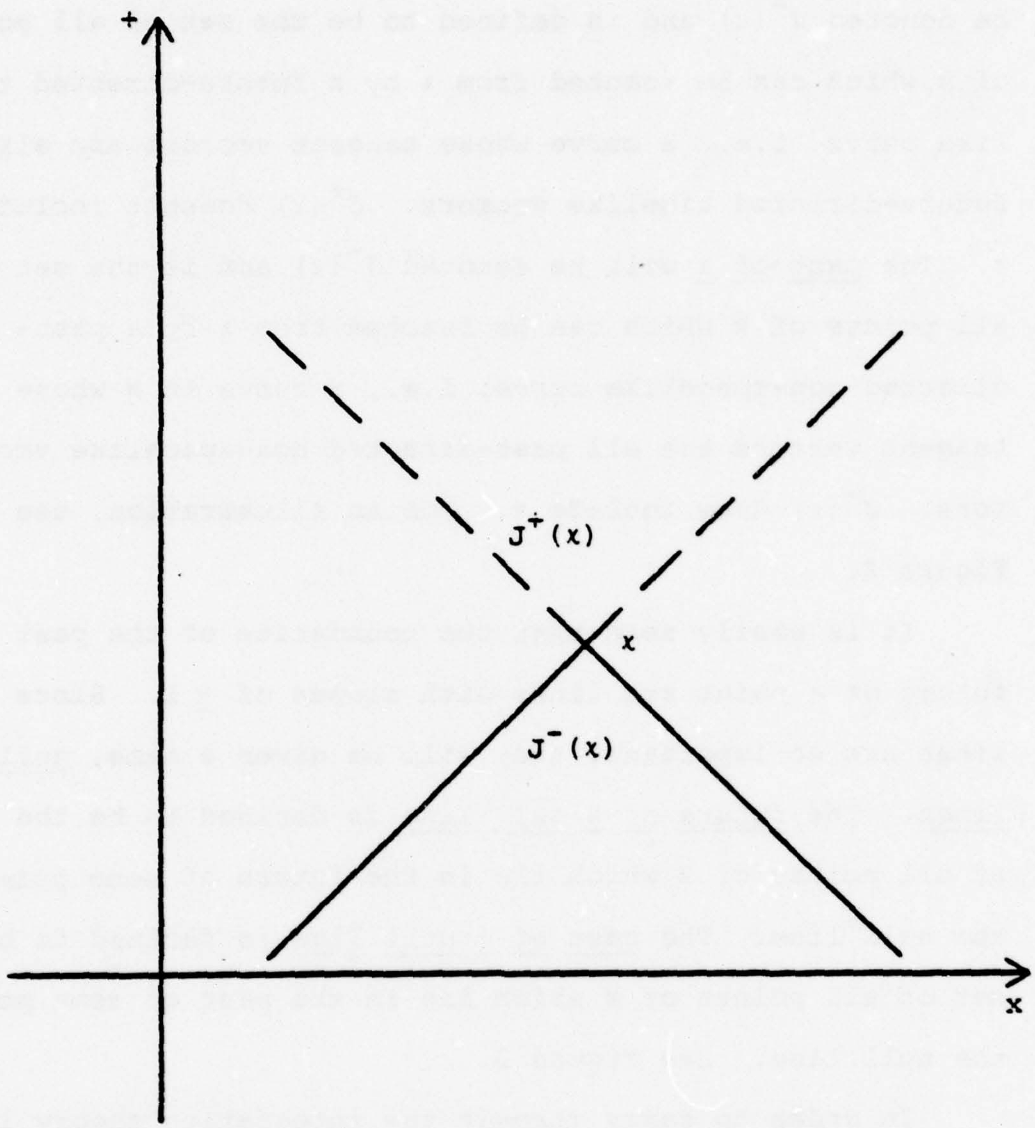
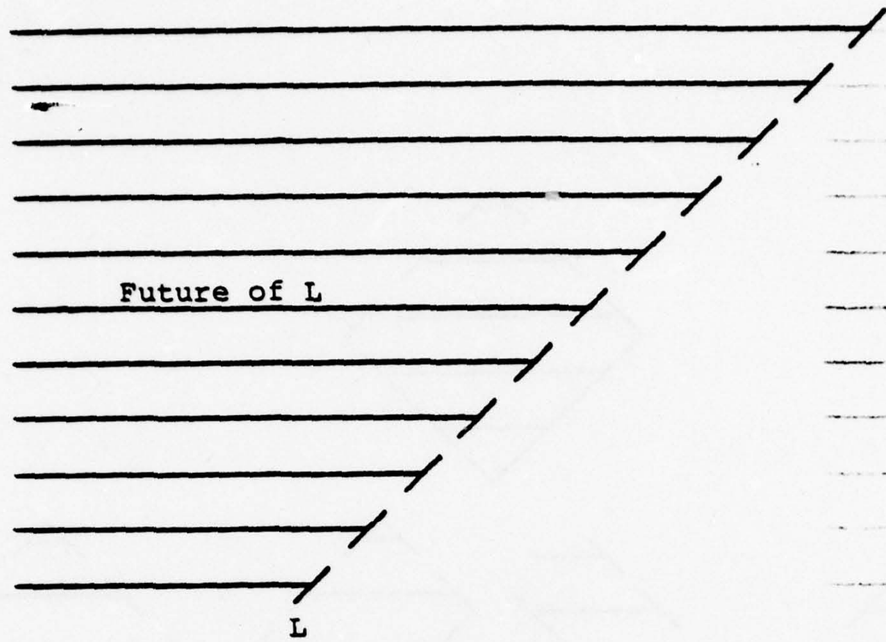


Figure 2.



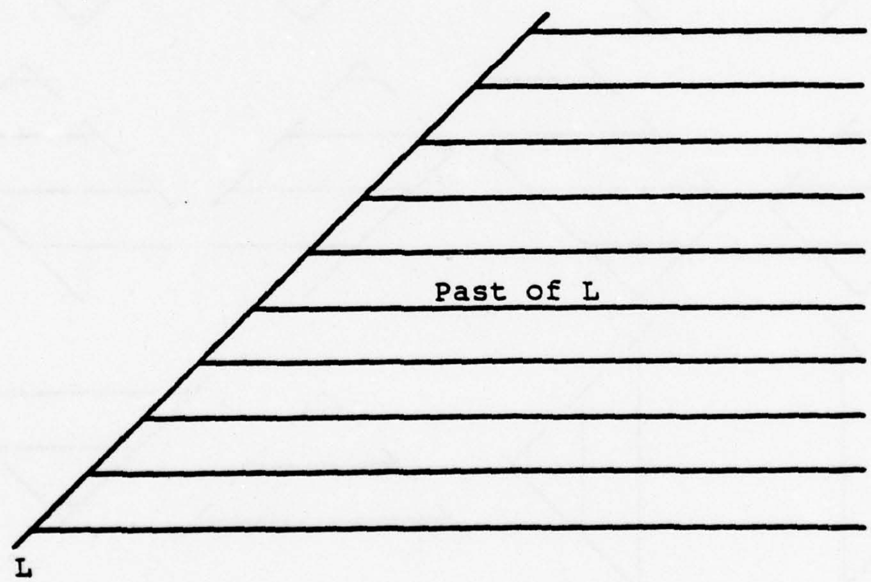


Figure 3

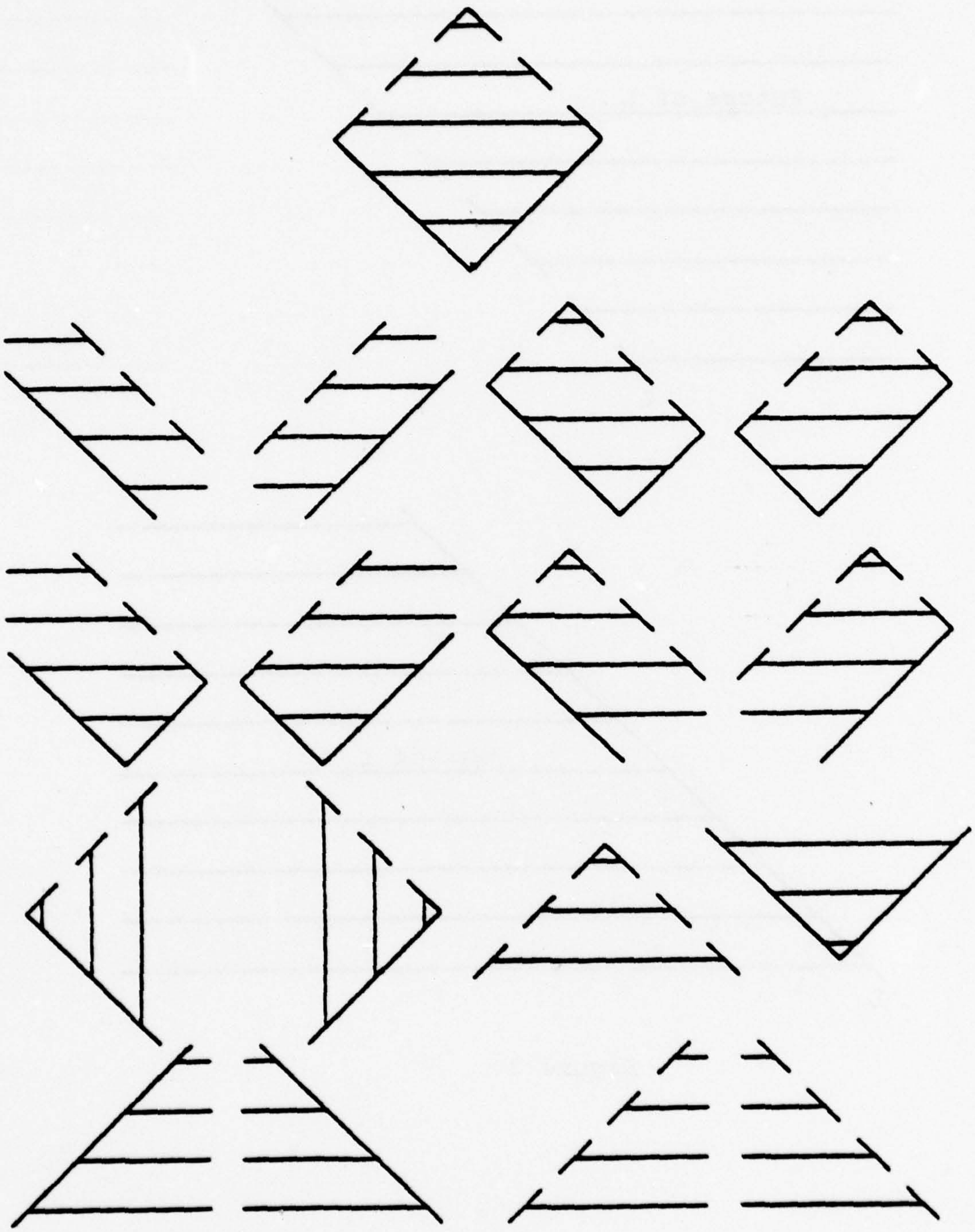


Figure 4.

the futures (pasts) of all the points contained in the diamond set and is denoted $J^+(D)$ [$J^-(D)$]. All possible diamond sets are shown in Figure 4.

Our first theorem concerns the suitability of the diamond sets for performing the integrations.

2.1 THEOREM The diamond sets form a semiring.

Proof: To establish this, we need to show that

i) the intersection of two diamond sets is a diamond set, and

ii) the set difference of two diamond sets is a disjoint union of diamond sets.

i) is immediately obvious from the definition of a diamond set. ii) is apparent from the diamond sets pictured in Figure 4 and the fact that all lines bounding the diamond sets have a slope of $\pm$ 1.

In order to keep the analogy with the classical case, we would like to be able to write the past and future of each diamond set as the past of a single point. With each null line L, associate two additional points u_L and ℓ_L. u_L will be called the upper point of L and has the property that the past of u_L is the past of L. The future of u_L will be the empty set. ℓ_L will be called the lower point of L and has the property that the future of ℓ_L is equal to the future of L. The past of ℓ_L is considered to be the empty set. We also add two more points, $+\infty$ and $-\infty$. The past of $+\infty$ and the future of $-\infty$ are M, and the future of $+\infty$ and the past of $-\infty$ are the empty set. These added points are purely a notational convenience, and as such, they have no relation to the points of M. We have not yet been able to discover whether or not there is a topology which would continuously extend the metric to include these points. With the inclusion of these extra points, it is now possible to write the past and future of any diamond set as the past and future of two unique points. For an example, see Figure 5.

Now let $L_2(M)$ be the Hilbert space of L_2 functions defined on M. If A is a subset of M, we define the projection

2.2. $$E(A): \; L_2(M) \to L_2(M)$$

by

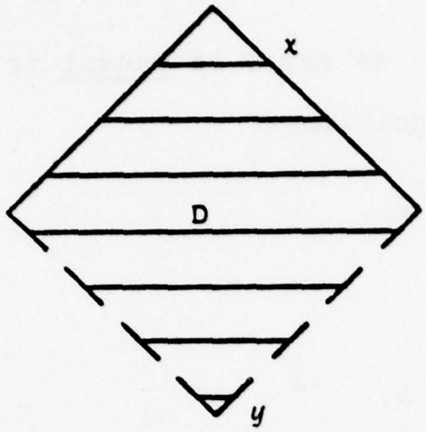
2.3. $$E(A)f(x) = \begin{cases} f(x), & x \in A \\ 0, & x \notin A \end{cases}, \quad f \in L_2(M) \; .$$

Analogously to the classical case, we define the projections E^x: $L_2(M) \to L_2(M)$ and E_x: $L_2(M) \to L_2(M)$ for $x \in M$ by

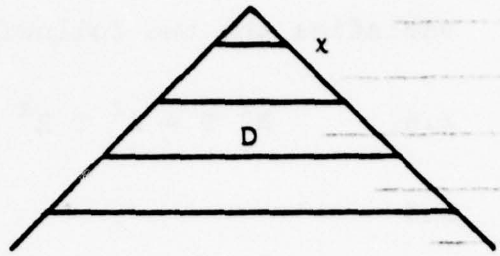
2.4. $$E^x = E[J^-(x)]$$

and

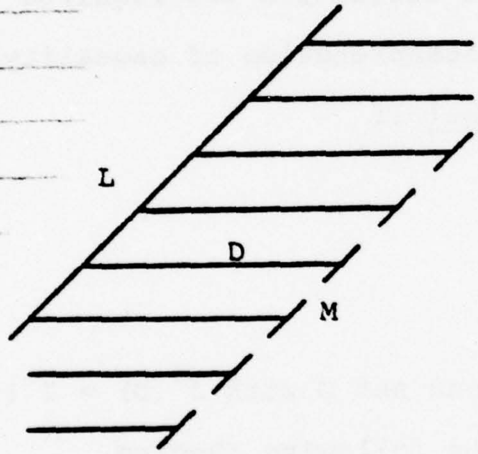
2.5. $$E_x = E[J^+(x)].$$



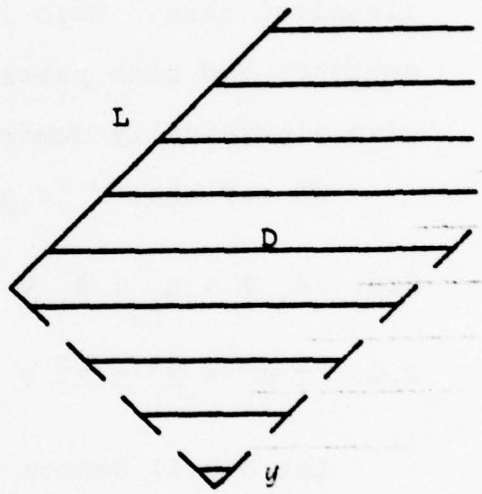
$J^+(D) = J^+(y)$
$J^-(D) = J^-(x)$



$J^+(D) = J^+(-\infty)$
$J^-(D) = J^-(x)$



$J^+(D) = J^+(\ell_M)$
$J^-(D) = J^-(u_L)$



$J^+(D) = J^+(y)$
$J^-(D) = J^-(u_L)$

Figure 5.

Let T be an operator on $L_2(M)$. We say T is <u>causal</u> if it satisfies the two following conditions:

2.6. $$E^x\, T = E^x\, T\, E^x \;\forall\, x \in M$$

and

2.7. $$T\, E_x = E_x\, T\, E_x \;\forall\, x \in M.$$

Notice that in the classical case. these two causality conditions imply each other. This is a result of the fact the ranges of E^x and E_x are complementary subspaces in the classical case. This is not the case in the relativistic setting, and both parts of the definition are required to give a physically desirable interpretation of causality.

We say that T is <u>anticausal</u> if

2.8. $$E_x\, T = E_x\, T\, E_x \;\forall\, x \in M$$

2.9. $$T\, E^x = E^x\, T\, E^x \;\forall\, x \in M.$$

Let $D(u,\ell)$ denote a diamond set D with $J^+(D) = J^+(\ell)$ and $J^-(D) = J^-(u)$. We have the following theorem.

2.2 <u>THEOREM</u> T: $L_2(M) \to L_2(M)$ is causal if and only if

2.10. $$E[D(u,\ell)]T = E[D(u,\ell)]T\, E^u$$

and

2.11. $$T\, E[D(u,\ell)] = E_\ell T\, E[D(u,\ell)]$$

for all diamond sets $D(u,\ell)$.

Proof: Suppose T is causal. Then $E^uT = E^uTE^u$ and $TE_\ell = E_\ell TE_\ell$. It is easily seen that $E[D(u,\ell)]E^u = [D(u,\ell)]$ and $E_\ell E[D(u,\ell)] = E[D(u,\ell)]$. Then $E[D(u,\ell)]E^uT = E[D(u,\ell)]E^uTE^u$ or $E[D(u,\ell)]T = E[D(u,\ell)]TE^u$. Similarly $TE[D(u,\ell)] = E_\ell TE[D(u,\ell)]$.

Now suppose $TE[D(u,\ell)] = E_\ell TE[D(u,\ell)]$ and $E[D(u,\ell)]T = E[D(u,\ell)]TE^u$ for all diamond sets. Let x be any point of M. Then $J^+(x)$ and $J^-(x)$ are diamond sets. It is also easily seen that $J^+[J^+(x)] = J^+(x)$ and $J^-[J^-(x)] = J^-(x)$. From this, it follows that $E^xT = E^xTE^x$ and $TE_x = E_xTE_x$,

We obtain a similar theorem for anticausal operators.

2.3 <u>THEOREM</u>: $T: L_2(M) \to L_2(M)$ is anticausal if and only if

2.12. $$TE[D(u,\ell)] = E^uTE[D(u,\ell)] ,$$

and

2.13. $$E[D(u,\ell)]T = E[D(u,\ell)]TE_\ell .$$

An operator $T: L_2(M) \to L_2(M)$ is said to be <u>memoryless</u> if T is both causal and anticausal. We have the following theorem.

2.4 <u>THEOREM</u>: T is memoryless if and only if

2.14. $$E[D(u,\ell)]T = TE[D(u,\ell)]$$

for every diamond set $D(u,\ell)$.

Proof: First assume T is memoryless. Then T is both causal and anticausal. T causal implies that

$$E[D(u,\ell)]T = E[D(u,\ell)]TE^{u} \text{ , and}$$

$$TE[D(u,\ell)] = E_{\ell}TE[D(u,\ell)],$$

and T anticausal implies that

$$E[D(u,\ell)]T = E[D(u,\ell)]TE_{\ell} \text{ , and}$$

$$TE[D(u,\ell)] = E^{u}TE[D(u,\ell)].$$

Then $E[D(u,\ell)]T = [E(D(u,\ell))E^{u} = [E(D(u,\ell))TE_{\ell}]E^{u}$

$$= E(D(u,\ell))TE(D(u,\ell)),$$

and

$$TE[D(u,\ell)] = E_{\ell}[TE(D(u,\ell))]$$

$$= E_{\ell}[E^{u}TE(D(u,\ell))]$$

$$= E(D(u,\ell))TE(D(u,\ell)).$$

Hence $E[D(u,\ell)]T = TE[D(u,\ell)]$ for all diamond sets $D(u,\ell)$.

Now suppose $E[D(u,\ell)]T = TE[D(u,\ell)]$ for all diamond sets $D(u,\ell)$.

Then $E[D(u,\ell)]T = E[D(u,\ell)]TE[D(u,\ell)]$

and $E[D(u,\ell)]TE^{u} = E[D(u,\ell)]TE[D(u,\ell)]E^{u}$

$$= E[D(u,\ell)]TE[D(u,\ell)] = E[D(u,\ell)]T.$$

Similarly $E_{\ell}TE[D(u,\ell)] = TE[D(u,\ell)]$. Hence T is causal. Similarly T is anticausal. Therefore T is memoryless.

The classes of operators which have been defined are as well-behaved as in the classical case.

2.5 THEOREM: The set of causal (anticausal, memoryless) bounded linear operators from a Banach algebra with identity which is closed in the strong operator topology of the algebra of all linear bounded operators.

Proof: The proof will be presented for the causal case only since the proof for the anticausal case is very similar, and the result for the memoryless case follows from the fact that the intersection of two Banach algebras is also a Banach algebra.

If T and S are causal, then

$$E^{x}TS = (E^{x}TE^{x})S = E^{x}T(E^{x}S) = E^{x}T(E^{x}SE^{x})$$
$$= (E^{x}TE^{x})SE^{x} = E^{x}TSE^{x}\ .$$

Similarly, $TSE_{x} = E_{x}TSE_{x}$, and
hence ST is causal. If we take the sum of S and T, we have $E^{x}(S+T) = E^{x}S + E^{x}T = E^{x}SE^{x} + E^{x}TE^{x} = E^{x}(S + T)E^{x}$, and $(S + T)E_{x} = E_{x}(S + T)E_{x}$.
Thus S + T is causal. Also, $E^{x}I = E^{x}E^{x}I = E^{x}IE^{x}$ and IE_{x}

$= E_x I E_x$. This proves the identity is causal. Finally, if T_i is a sequence of causal operators converging strongly to T, i.e., $\lim_{i\to\infty} T_i f = Tf$ for all f, then since E^x is bounded,

$$E^x Tf = E^x[\lim_{i\to\infty} T_i f] = \lim_{i\to\infty} E^x T_i f$$

$$= \lim_{i\to\infty} E^x T_i E^x f = E^x[\lim_{i\to\infty} T_i E^x f] = E^x T E^x f.$$

Similarly, $TE_x f = E_x T E_x f$ for all f. Hence T is causal.

We have the following theorem for the adjoint of a causal operator.

2.6 <u>THEOREM</u>: An operator T is causal if and only if T* is anticausal.

Proof: Suppose T is causal. Then $E^x T = E^x T E^x$. Taking the adjoint of both sides, we have $(E^x T)^* = (E^x T E^x)^*$ which reduces to $T^* E^x = E^x T^* E^x$. Similarly, $E_x T^* = E_x T^* E_x$. Hence T* is anticausal. The converse is similar.

Since a memoryless operator is both causal and anticausal, then so is its adjoint and hence the adjoint of a memoryless operator is also memoryless.

At the moment, nothing can be said about the inverse of a causal operator in the relativistic case. The results from the classical case don't carry over, partly because the ranges of the projections aren't complementary subspaces, and partly because the condition $E^x f = E^x g \Rightarrow E^x Tf =$

$E^{x}Tg$, $\forall\ x \in M$, $f,g \in L_2(M)$, is not enough to insure causality for T in the relativistic case.

We come now to the extension of the integrals of triangular truncation to the relativistic setting. It was for this purpose that the diamond sets were introduced. Since they form a semiring, they can be used to partition M so that the integrals can be defined [5]. The upper Cauchy integral of an operator valued function f on M is defined by

$$2.15. \quad UC \int f(x)dE(x) = \lim_{p \in P} \sum_{i=1}^{n(p)} f(u_i)E[D_i(u_i,\ell_i)]$$

where the limit exists in the uniform topology over the net of all partitions of M into diamond sets $D_i(u_i,\ell_i)$. Similarly, the lower Cauchy integral is defined by

$$2.16. \quad LC \int f(x)dE(x) = \lim_{p \in P} \sum_{i=1}^{n(p)} f(\ell_i)E[D_i(u_i,\ell_i)].$$

These integrals can also be defined with the measure on the left, or over a portion of M instead of all M. We can also define the strong Cauchy integrals $SUC\int$ and $SLC\int$ by taking the limit in the strong operator topology. We then have the following theorem relating causality to these Cauchy integrals.

2.7 THEOREM: The following are equivalent for a linear bounded operator T.

i) T is causal.

2.17. ii) $UC \int dE(x)TE^{x} = LC \int E_{x}TdE(x) = T.$

2.18. iii) $SUC \int dE(x)TE^{x} = SLC \int E_{x}TdE(x) = T.$

Proof: i) => ii).

If T is causal, then $E[D(u,\ell)]TE^{x} = E[D(u,\ell)]T$ and $E_{x}TE[D(u,\ell)] = TE[D(u,\ell)]$ for all $x \in M$. Hence for any partition of M into diamond sets,

$$T = IT = [\lim_{p\in P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]]T$$

$$= \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]T$$

$$= \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]TE^{u_i}$$

$$= UC \int dE(x)TE^{x}.$$

Similarly, $T = LC \int E_{x}TdE(x)$.

ii) => iii).

This follows immediately from the fact that uniform convergence implies strong convergence.

iii) => i.

We have $SUC \int dE(x)TE^{x} = SLC \int E_{x}TdE(x) = T$. Then $E^{y}T = E^{y}SUC\int_{M}dE(x)TE^{x} = SUC\int_{J^{-}(y)}dE(x)TE^{x} = SUC\int_{J^{-}(y)}dE(x)TE^{x}E^{y}$

$= [SUC\int_{J^-(y)} dE(x)TE^x]E^y. = E^yTE^y$. Similarly, $TE_y = E_yTE_y$. Hence T is causal.

A similar theorem is also true for anticausal operators.

2.8 <u>THOEREM</u>: The following are equivalent for a bounded linear operator T.

i) T is anticausal.

2.19. ii) $UC\int E^x TdE(x) = LC\int dE(x)TE_x = T$.

2.20. iii) $SUC\int E^x TdE(x) = SLC\int dE(x)TE_x = T$.

For an integral representation of memoryless operators, we need to define the diagonal Cauchy integral

2.21. $$C\int dE(x)TdE(x) = \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]TE[D_i(u_i,\ell_i)]$$

where the limit is taken in the uniform topology over the net of all partitions of M into diamond sets. We can also define the strong diagonal Cauchy integral $SC\int dE(x)TdE(x)$ by taking the limit in the strong topology. We then have the following theorem for memoryless operators.

2.9 <u>THEOREM</u>: For a bounded linear operator T, the following are equivalent.

i) T is memoryless.

2.22. ii) $C\int dE(x)TdE(x) = T$.

2.23. iii) $SC\int dE(x)TdE(x) = T$.

Proof: Similar to the proof of the theorem for causal operators.

We now define the notion of strict causality. A bounded linear operator T is said to be strictly causal if

2.24. $LC\int dE(x)TE^{x} = UC\int E_{x}TdE(x) = T$.

A bounded linear operator is said to be strongly strictly causal if

2.25. $SLC\int dE(x)TE^{x} = SUC\int E_{x}TdE(x) = T$.

The strictly causal case is different from the causal case in that strict causality implies strong strict causality but not conversely (see [1]).

In order to characterize the relationship between the strictly causal and causal operators, we need to define the following integrals.

2.26. $\underline{S}\int dE(x)T\underline{E}^{x} = \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]T(E^{u_i} - E^{\ell_i})$.

2.27. $\overline{S}\int \overline{E}_{x}TdE(x) = \lim_{p\in P} \sum_{i=1}^{n(p)} (E_{\ell_i} - E_{u_i})TE[D_i(u_i,\ell_i)]$.

The limit is taken in the uniform topology over the net of all partitions of M into diamond sets. We can also define the integrals

$$S\underline{S}\int dE(x)T\underline{E}^{x} \text{ and } S\overline{S}\int \overline{E}_{x}TdE(x)$$

where the l[illegible]t is taken in the strong topology. We then have the following theorem for strictly causal operators.

2.10 <u>THEOREM</u>: A bounded linear operator T is strictly causal if and only if T is causal and

$$2.28. \quad \underline{S}\int dE(x)T\underline{E}^{x} = \overline{S}\int \overline{E}_{x}TdE(x) \ .$$

Proof: First suppose that T is strictly causal.

$$\begin{aligned} E^{x}T &= E^{x}LC\int_{M}dE(y)TE^{y} = LC\int_{J^{-}(x)}dE(y)TE^{y} \\ &= LC\int_{J^{-}(x)}dE(y)TE^{y}E^{x} = E^{x}TE^{x}. \end{aligned}$$

Similarly $TE_{x} = E_{x}TE_{x}$. Hence T is causal. T being causal implies that

$$\begin{aligned} T &= UC\int dE(x)TE^{x} \\ &= \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_{i}(u_{i},\ell_{i})]TE^{u_{i}} \\ &= \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_{i}(u_{i},\ell_{i})]T(E^{\ell_{i}} + [E^{u_{i}} - E^{\ell_{i}}]) \\ &= \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_{i}(u_{i},\ell_{i})]TE^{\ell_{i}} + \lim_{p\in P} \sum_{i=1}^{n(p)} E[D_{i}(u_{i},\ell_{i})]T \\ &\qquad [E^{u_{i}}-E^{\ell_{i}}] \\ &= LC\int dE(x)TE^{x} + \underline{S}\int dE(x)T\underline{E}^{x} \\ &= T + \underline{S}\int dE(x)T\underline{E}^{x}. \end{aligned}$$

Hence $\underline{S}\int dE(x)T\underline{E}^x = 0$. Similarly $\overline{S}\int\underline{E}_x TdE(x) = 0$.

By reversing the above argument, it can be seen that T causal and $\underline{S}\int dE(x)T\underline{E}^x = \overline{S}\int\overline{E}_x TdE(x) = 0$ implies that T is strictly causal.

We have a similar theorem for strongly strictly causal operators.

2.11 <u>THEOREM</u>: A bounded linear operator T is strongly strictly causal if and only if T is causal and

$$2.29. \quad S\underline{S}\int dE(x)T\underline{E}^x = S\overline{S}\int\overline{E}_x TdE(x) = 0.$$

There are also similar theorems for strictly anticausal and strongly strictly anticausal operators. We first need to define the following integrals.

$$2.30. \quad \underline{S}\int\underline{E}^x TdE(x) = \lim_{p\epsilon P}\sum_{i=1}^{n(p)}[E^{u_i}-E^{\ell_i}]TE[D_i(u_i,\ell_i)].$$

$$2.31. \quad \overline{S}\int dE(x)T\underline{E}_x = \lim_{p\epsilon P}\sum_{i=1}^{n(p)}E[D_i(u_i,\ell_i)]T[E_{\ell_i}-E_{u_i}].$$

Again, the limit is taken in the uniform topology over the net of all partitions of M into diamond sets. Then we have the following theorem which is proved in the same manner as the corresponding theorem on strictly causal operators.

2.12 <u>THEOREM</u>: A bounded linear operator T is strictly anticausal if and only if T is anticausal and

2.32. $\underline{S}\int \underline{E}^{x} T dE(x) = \overline{S}\int dE(x) T \underline{E}_{x} = 0.$

If we take limits in the strong topology, we obtain the following theorem for strongly strictly anticausal operators.

2.13 <u>THEOREM</u>: A bounded linear operator T is strongly strictly anticausal if and only if T is anticausal and

2.33. $S\underline{S}\int \underline{E}^{x} T dE(x) = S\overline{S}\int dE(x) T \underline{E}_{x} = 0.$

We have the following theorem for the space of strictly causal (strictly anticausal, strongly strictly causal, strongly strictly anticausal) operators.

2.14 <u>THEOREM</u>: The space of strictly causal (strictly anticausal, strongly strictly causal, strongly strictly anticausal) operators forms a Banach space which is closed in the uniform operator topology of the space of all bounded linear operators.

Proof: The proofs in all four cases are similar, so only the strictly causal case will be presented.

These operators form a Banach space since they are defined by a linear equation $T = LC\int dE(x) T E^{x}$. Now suppose $T_i \to T$ where the T_i are strictly causal. Since the T_i are strictly causal, they are causal, and hence T is causal. We would now like to show that $\underline{S}\int dE(x) T E^{x} = 0$.

For any $\varepsilon > 0$, choose j such that $||T_j - T|| < \varepsilon/2$ and a partition of M into diamond sets $D_i(u_i,\ell_i)$, $i = 1,\ldots,n$ such that

$$||\sum_{i=1}^{n} E[D_i(u_i,\ell_i)]T_j(E^{u_i})|| < \varepsilon/2.$$

Then $||\sum_{i=1}^{n} E[D_i(u_i,\ell_i)]T(E^{u_i}-E^{\ell_i})|| = ||\sum_{i=1}^{n}E[D_i(u_i,\ell_i)](T-T_j)$ $(E^{u_i}-E^{\ell_i}) + \sum_{i=1}^{n}E[D_i(u_i,\ell_i)]T_j(E^{u_i}-E^{\ell_i})|| \leq \sup_i||E[D_i(u_i,\ell_i)]$ $(T-T_j)(E^{u_i}-E^{\ell_i})|| + \sup_i||E[D_i(u_i,\ell_i)]T_j(E^{u_i}-E^{\ell_i})|| < \varepsilon/2 +$ $\varepsilon/2 = \varepsilon$. So the partial sums for $S\int dE(x)TE^x$ converge to zero. Similarly, $\overline{S}\int\overline{E}_x TdE(x) = 0$. Hence T is strictly causal.

The following theorem relates the strictly causal and the strictly anticausal operators to each other.

2.15 <u>THEOREM</u>: *An operator* T *is (strongly) strictly causal if and only if* T* *is (strongly) strictly anticausal.*

Proof: $LC\int dE(x)TE^x = \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)TE^{\ell_i}$. If we take the adjoint of $\sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]TE^{\ell_i}$, then we have $\sum_{i=1}^{n(p)} E^{\ell_i}T^*E[D_i(u_i,\ell_i)]$. Since the adjoint is a linear isometry in the space of bounded linear operators on a Hilbert space, the first integral converges to T if and only if the second integral converges to T*. Using a similar procedure for $UC\int E_x\, TdE(x)$, we see that T is strictly causal if and only if T* is strictly anticausal.

In order to state and prove the additive decomposition theorem for arbitrary operators, we first need to define a new class of operators which will be called spacelike. For a diamond set $D(u,\ell)$ we define the projection

2.34. $$E_u^\ell = I - E^\ell - E_u.$$

With this definition, we say that an operator T is <u>spacelike</u> if

2.35. $$TE[D(u,\ell)] = E_u^\ell TE[D(u,\ell)]$$

and

2.36. $$E[D(u,\ell)]T = E[D(u,\ell)]TE_u^\ell$$

for all diamond sets $D(u,\ell)$.

We can obtain an integral characterization of spacelike operators by defining the following integrals.

2.37. $$S\int E_x^x T dE(x) = \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E_{u_i}^{\ell_i} TE[D_i(u_i,\ell_i)].$$

2.38. $$S\int dE(x) TE_x^x = \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E[D_i(u_i,\ell_i)]TE_{u_i}^{\ell_i}.$$

The limit is taken in the uniform topology over the set of all partitions of M into diamond sets. We can also take limits in the strong topology in which case we obtain the integrals $SS\int E_x^x T dE(x)$ and $SS\int dE(x) TE_x^x$. We then have the following theorem.

2.16 <u>THEOREM</u>: For a bounded linear operator T, the following are equivalent.

i) T is spacelike.

2.39. ii) $S\int E_x^x T dE(x) = S\int dE(x)\, TE_x^x = T$.

2.40. iii) $SS\int E_x^x T dE(x) = SS\int dE(x) TE_x^x = T$.

Proof: Similar to the proof for the causal case.

We also obtain the following theorem for the relationship between a spacelike operator T and its adjoint T*.

2.17 <u>THEOREM</u>: T is spacelike if and only if T* is spacelike.

Proof: $TE[D(u,\ell)] = E_u^\ell TE[D(u,\ell)] \Leftrightarrow E[D(u,\ell)]T^*$

$$= E[D(u,\ell)]T^*E_u^\ell \quad \text{and,}$$

$$E[D(u,\ell)]T = E[D(u,\ell)]TE_u^\ell \Leftrightarrow T^*E[D(u,\ell)]$$

$$= E_u^\ell T^*E[D(u,\ell)].$$

Finally, we obtain the additive decomposition theorem for an arbitrary bounded linear operator.

2.19 <u>THEOREM</u>: Let T be an arbitrary bounded linear operator. Then T can be decomposed as T = C + A + S where C is strictly causal, A is strictly anticausal, and S is space-like if and only if the integrals

$$UC\int E_x T dE(x), \quad LC\int dE(x) T E^x,$$

$$LC\int E^x T dE(x), \quad UC\int dE(x) T E_x,$$

$$S\int E_x^x T dE(x) \text{ and } S\int dE(x) T E_x^x$$

exist and satisfy

2.41. $UC\int E_x T dE(x) = LC\int dE(x) T E^x$,

2.42. $LC\int E^x T dE(x) = UC\int dE(x) T E_x$,

and

2.43. $S\int E_x^x T dE(x) = S\int dE(x) T E_x^x$,

in which case the decomposition is given by

2.44. $C = UC\int E_x T dE(x) = LC\int dE(x) T E^x$,

2.45. $A = LC\int E^x T dE(x) = UC\int dE(x) T E_x$,

and

2.46. $S = S\int E_x^x T dE(x) = S\int dE(x) T E_x^x$.

Proof: First assume that all the integrals exist and satisfy the required relationships. Then

$$T = UC\int TdE(x) = \lim_{p\in P} \sum_{i=1}^{n(p)} TE[D_i(u_i,\ell_i)]$$

$$= \lim_{p\in P} \sum_{i=1}^{n(p)} [E^{\ell_i} + E_{u_i} + E_{u_i}^{\ell_i}]TE[D_i(u_i\ell_i)]$$

$$= \lim_{p\in P} \sum_{i=1}^{n(p)} E^{\ell_i}TE[D_i(u_i,\ell_i)] + \lim_{p\in P} \sum_{i=1}^{n(p)} TE[D_i(u_i,\ell_i)]$$

$$+ \lim_{p\in P} \sum_{i=1}^{n(p)} E_{u_i}^{\ell_i}TE[D_i(u_i,\ell_i)]$$

$$= LC\int E^x TdE(x) + UC\int E_x TdE(x) + S\int E_x^x TdE(x).$$

Similarly,

$$T = LC\int dE(x)T$$

$$= UC\int dE(x)TE_x + LC\int dE(x)TE^x + S\int dE(x)TE_x^x$$

We have

$$UC\int E_x CdE(x) = UC\int E_x[UC\int E_y TdE(y)]dE(x)$$

$$= UC\int E_z TdE(z) = C.$$

Similarly, $LC\int dE(x)CE^x = C$.

Hence C is strictly causal.

Similarly, A is strictly anticausal, S is space-like, and therefore this is the desired decomposition.

Now, suppose that the desired decomposition exists. Then

$$UC\int E_x T\,dE(x) = UC\int E_x (C+A+S)\,dE(x)$$

$$= UC\int E_x C\,dE(x) + UC\int E_x A\,dE(x) + UC\int E_x S\,dE(x).$$

$$UC\int E_x A\,dE(x) = \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E_{u_i} AE[D_i(u_i, \ell_i)]$$

$$= \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E_{u_i} E^{u_i} AE[D_i(u_i, \ell_i)] = 0.$$

$$UC\int E_x S\,dE(x) = \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E_{u_i} SE[D_i(u_i. \ell_i)]$$

$$= \lim_{p\varepsilon P} \sum_{i=1}^{n(p)} E_{u_i} E_{u_i}^{\ell_i} SE[D_i(u_i, \ell_i)] = 0.$$

Hence $UC\int E_x T\,dE(x) = UC\int E_x C\,dE(x) = C$.
Similarly $LC\int dE(x) TE^x = C$. Thus C is of the required form. In the same manner, A and S can be shown to have the required form.

We have a similar theorem for additive decomposition into operators defined by strongly convergent integrals.

2.20 <u>THEOREM</u>: Let T be an arbitrary linear bounded operator. Then T can be decomposed as $T = C + A + S$ where C is strongly strictly causal, A is strongly strictly anticausal, and S is spacelike if and only if the integrals

$$SUC\int E_x T dE(x) \quad , \quad SLC\int dE(x) T E^x,$$

$$SLC\int E^x T dE(x) \quad , \quad SUC\int dE(x) T E_x,$$

$$SS\int E_x^x T dE(x) \quad \text{and} \quad SS\int dE(x) T E_x^x$$

exist and satisfy

2.47. $SUC\int E_x T dE(x) = SLC\int dE(x) T E^x$,

2.48. $SLC\int E^x T dE(x) = SUC\int dE(x) T E_x$,

and

2.49. $SS\int E_x^x T dE(x) = SS\int dE(x) T E_x^x$,

in which case the decomposition is given by

2.50. $C = SUC\int E_x T dE(x) = SLC\int dE(x) T E^x$,

2.51. $A = SLC\int E^x T dE(x) = SUC\int dE(x) T E_x$,

and

2.52. $S = SS\int E_x^x T dE(x) = SS\int dE(x) T E_x^x$.

CHAPTER III

CONCLUSIONS

Following the development of resolution space for the two-dimensional special relativistic case, it would be nice to extend the development to the four-dimensional special relativistic case, and then to the general relativistic case. However, no way has been found to extend the results in this paper to even the three-dimensional special relativistic setting.

One major problem has been the attempt to keep a point development of resolution space. The lack of an ordering on the space-time manifolds presented difficulties even in the two-dimensional case. By using diamond sets it was possible to keep a semblance of the point development, but even in this case special points had to be introduced, and attempts to extend the space-time topology to include the extra points in such a manner as to continuously extend the metric were fruitless.

We attempted to extend the diamond sets to higher dimensions, but we were unable to find a higher dimensional analogue of the diamond sets. Hence, it would appear that the diamond sets are peculiar to two dimensions. However, the Borel sets in two dimensions can be generated from the diamond sets, and possibly a higher dimensional development could proceed from a purely set viewpoint without recourse to points. Such a development would use the spectral measure E with its projections E(A) and not even try to bother with point projections such as E^x. We feel that this technique would end up producing much the same results paralleling classical resolution space as in the two-dimensional case.

Although most of the results mirrored those of classical resolution space, one new type of operator was obtained which appears to be a generalization of the memoryless operator. This is the spacelike operator which was introduced in the decomposition theorem. It is conjectured that if the inverse of a spacelike operator exists, then it will also be spacelike.

However, other than this one interesting conjecture the extension of resolution space to a relativistic setting seems to offer little promise of new insights into operator theoretic systems theory.

REFERENCES

[1] R. Saeks, Resolution Space, Operators and Systems, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[2] C. W. Misner, K. S. Thorne, and J. A. Wheeler, Gravitation, Freeman, San Francisco, 1973.

[3] S. W. Hawking, and G. F. R. Ellis, The Large Scale Structure of Space-Time, Cambridge, London, New York, 1973.

[4] J. D. Jackson, Classical Electrodynamics, 2nd edition, Wiley, New York, London, 1975.

[5] R. Saeks, and R. A. Goldstein, Indiana University Mathematics Journal, 4 (1972), 367.

[6] R. Penrose, Techniques of Differential Topology in Relativity, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1972.